War games and Honeypots: the role of role-playing in assessment

Helen Ashman[1], Wayne Gartner[1] and Kirsten Wahlstrom[1],

[1] School of Computer and Information Science
University of South Australia
{helen.ashman, garwd001, kirsten.wahlstrom}@unisa.edu.au

Abstract. The assessment of large classes of students is generally time-consuming for the lecturer as well as somewhat impersonal for the students. The limits of the lecturer's time available for feedback is the main constraint with larger classes, inevitably leading to less time per student and a concomitant lower level of feedback. Engaging the students' interest can also be challenging in a larger class, with many students feeling disenfranchised by the lack of individual attention. This paper discusses the use of a "War Games" scenario in assessing students of Computer and Network Security with group-based and problem-based learning exercises. The students contribute directly to creating the environment of the exercise, assessing their peers at two levels and creating part of the examinable materials as well as exchanging mutual feedback in a post-mortem session. This sharing of the assessment duties amongst the student body supplements and augments the lecturer's own grading. The students gain in many ways, including more detailed, comprehensive and timely feedback from their peers. The War Games assessed exercise is a popular and highly instructive exercise, but requires caution in its setting up, since it is essential that the activities of the exercise be contained within the controlled environment.

Keywords: war games, experiential learning

1 Introduction

1.1 Motivation

Engaging students in their own education can be a challenge, and teachers are under pressure to ensure students qualify with their maximum possible outcomes. However, communicating enthusiasm for their assessed exercises to students can still be a challenge, yet enthusiasm is a strong motivator for effort and thus the best possible result for a student. Teachers need to exploit any feature that makes students enthusiastic about their work, and motivates them to put in at least the requisite time.

Problem-based learning and group-based learning aim can ease the assimilation of information and skills by students, even (or especially) those who are not academically strong. Problem-based, or experiential, learning is increasingly favoured as a “doing” teaching mechanism that replaces or supplements the conventional

formal lecturing style. It involves students in hands-on, problem-solving exercises that mimic the possible work styles they will use in later life.

Group-based learning encourages students to work in teams and can be associated with peer assessment/review of performance at the individual level. There can also be scope for the groups to interact with each other as a core part of the exercise, enabling peer review not just at the individual level but also at the group level.

The two assessments described here are both problem-based and group-based. A major feature is that students are involved in significant role-playing that reflects real-life situations. They reflect the possible workplace that the students may find themselves in, in terms of both the work they will be doing and their work team composition. In each case, students are working in an adversarial, "war games" situation, using each other as adversaries in one and real adversaries in the other.

Both are role-playing games. There is no story-telling as such, nor any need for it. They are, however, educational games, whose purpose is to teach the students about computer and network security and computer forensics principles.

1.2 Types of group-based assessment

In some assessed exercises, the students gain an excellent understanding of the performance of their peers through interactions forming part of the exercise. The most common form of these interactions is working in groups on a common project. However there are other forms of interaction which can contribute to assessment.

Security of computers and networks is a fundamentally adversarial technology, with the “good guys” being network defenders (system administrators) and the “bad guys” being hackers. In the exercises described in this paper, students play both roles. There are different styles of group-based exercise, namely:

- *independent* – each group operates generally without interaction with other groups
- *competitive* – groups compete to produce the best/fastest/etc. results, not interacting with each other but measuring their performance against each other
- *co-operative* – groups work collaboratively to contribute to a larger project
- *adversarial* – groups interact negatively with each other to degrade each other’s work, while simultaneously defending their own work from others.

The adversarial style requires the construction of a scenario where adversarial roles are both feasible and meaningful. A simple example is to divide a sporting team into two distinct teams and pit them against each other. Likewise debating teams practise against each other, and military academies train staff in adversarial teams. In academia, scenarios that admit adversarial exercises generally mimic adversarial situations in the real world. Politics and finance are possibilities, while the "war games" exercise directly reflects the adversarial nature of computer security.

This paper discusses one example of the adversarial style of exercise. The "war games" in section 2 is an adversarial group-based exercise while the "honeypot" in section 3 is an independent group-based exercise, although one which involves potential contact with real adversaries when the students release their websites on the Internet. Section 4 discusses the variety and benefits of the roles played and how the war games exercise meets experiential learning specifications.

1.3 Peer review

Peer review is often used when assessing group exercises, acknowledging the students' own insights into the activity and participation of their colleagues. In the "war games" exercise, peer review is extended to the group level as well. The exercise is at its core about managing the interaction between groups, with each group endeavouring to prevent hostile actions from other groups while servicing authorised actions, usually from the same groups. At the same time, each group is themselves perpetrating both types of actions on other groups. This gives each group excellent exposure to the quality of the work of other groups, allowing them to accurately quantify the performance of every other group.

Peer review, at both levels, is a useful tool for aiding students in understanding the quality of their own achievements and that of their group. It is generally accepted that the less competent a student is, the more they may overinflate their performance, having little idea of how poorly they compare to others in their cohort [7].

Aside from the undoubted benefit of understanding one's personal and group performance within the cohort, peer review serves two other useful functions:

peer review generates a realistic environment for the exercise: each group must assess the performance and availability of the services offered by other groups at least twice weekly, by attempting to use these services (such as file transfer and Web pages, including the Visitor Book) in a non-hostile fashion, as if they were a normal user. The actions that the groups perform to do this group-level peer assessment are what comprise the "normal" traffic in the environment, and without these actions there would be significantly less normal traffic and the proportion of hostile traffic would be unrealistically high. Thus the traffic generated by regular peer review is essential to the realism of the environment.

peer review saves the lecturer effort in marking: without the regular peer review between groups, the course convenor would have to perform these assessments to establish how well each group achieved the level of services required. Peer review devolves this activity to the groups. It would be unsupportably labour-intensive for the course convenor to provide a similar level of assessment. So while the course convenor must still mark reports, they need not assess each group's services daily.

peer review provides detailed assessment that would otherwise not be feasible: far more detail in the assessment of groups is possible through the peer review at group level, partly because of the much larger pool of effort that can go into the assessment - between 12 and 20 groups will be reviewing each other, providing a parallel and more frequent assessment of each others' services.

It should also be noted that the groups perpetrating an attack are in the best possible position to assess the impact on the targeted group, having an insider's view of the perpetrated attacks. Other groups may notice that a targeted group is not providing services, but the attacking group will know exactly what happened and what weaknesses in the other group's defences allowed it to succeed.

Peer review does carry penalties, such as the need to preserve the confidentiality of the peer review means that some transparency of marking is sacrificed.

In summary, the peer review process is a useful tool at both individual and group levels, but is also an essential component of the operational environment of the assessed exercise.

2 The War Games Exercise

2.1 Overview

The war games exercise is an adversarial group-based exercise which is in its eleventh year in 2009, having run for eight years at the University of Nottingham and at the University of South Australia since 2007. Some features are reported in prior publications [1] [2] and the original course details are available online [3] but a brief recapitulation of its operations are given in this section.

The pertinent features of the war games exercise for this paper are

- role playing at many levels;
- two levels of peer assessment: group and individual;
- realism of the exercise to real life, not just as a potential job, but in everyday life as a user of computers.

The war games exercise forms a significant part of the teaching and assessment of *Computer and Network Security* at UniSA. It is both problem-based and group-based, with the groups interacting strongly within an adversarial scenario. In fact, the interactions are an essential part of the exercise as they provide the environment for testing and assessing the groups' work and create a whole-class learning environment that encourages greater student participation through peer involvement.

The environment is a major assessed component of the course, and is based on having the students themselves create the context in which their exercise operates. Students are collected into groups and each group is assigned the task of setting up Internet services for use by all the other groups in the class. After setting up their services and securing them as they think best, students then have the triple technical role of attacking ("hacking") the services belonging to all other groups while simultaneously defending their own services against other attackers, along with the third role of making "normal" (non-attacking) use of the services provided by the other groups, during which time they assess each other group's defensive performance, as well as providing a background of normal traffic in their network.

All three roles are essential to the successful operation of the exercise:

- *the defending role* constitutes the main expertise to be gained from the exercise
- *the hacking role* is needed to give the defending roles something to work with.
- *the normal user role* is essential for two reasons. Firstly it generates a quantity of "normal" or non-hostile traffic without which all traffic would necessarily be hostile. Part of the defending role is to determine which traffic is hostile, while still normally servicing non-hostile requests. Secondly the normal role provides an accurate way to assess the performance of other groups by regularly testing the required services of the defending roles.

The hacking role has the additional benefit of giving the defending role insights into the mindset of the attacker, in a "poacher turned gamekeeper" situation.

The assessed elements of this course are 40% on an exam, and the remaining 60% on the war games exercise. 20% of that 60% comes from the group report, which is the same for each member of the group unless a particularly low level of participation is evident. A further 10% for each is averaged from the marks given by each other group for that group's performance in resisting or recovering from hacking attacks,

again with the same mark shared by all group members. There are two individual components, with each student handing in an individual report worth 10% describing and rating their contribution to the group's work, and a further 10% is averaged from peer review by other members of the student's group. Thus the students essentially mark the performance of each individual and each group through the peer review and group review components.

A final 10% is given to each group, again shared by all members, based on a presentation in a seminar session which takes the place of one of the lectures. Each group presents its activities and its analysis of what actually happened during the exercise. This post-mortem seminar session is an essential part of the knowledge gaining process as well as providing timely feedback, since the adversarial nature of the war game means that students must keep their activities secret while the exercise is running - the other groups must deduce what they can about the hostile actions in order to deal with them. A post-mortem at the conclusion of the exercise allows them to understand what happened and gave them feedback about their real-time analyses.

2.2 Technical specifications

This section details the network environment to capture and analyse network traffic for the war games exercise.

There are some basic hardware requirements for the exercise:

- *Laboratory with computers*: one previous presentation of the course ran the exercise using virtual machine software, but this was not ideal since students had to learn how to use the virtual machine software as an additional learning overhead, during an already intensive exercise. This also added to the workload of support staff. Thus we now insist on using real computers in a laboratory, especially since most students already have adequate background in the use and to some extent, the setting up of hardware, operating systems and some services.
- *Switch*: The switch is a key component of the network and is the means of connecting the computers so they can communicate with each other. The switch is also the means of mirroring network traffic for capturing and further analysis.
  In a previous year of building the network, a hub was used and not a switch. However a hub operates insecurely by forwarding traffic out on every port, while a switch will only send traffic on required ports, much more secure for the participating machines. However, this now means the analysing machine can no longer access all traffic. If we were to put the Sniffer onto the switch as is, the sniffer would only receive:
    - Traffic destined to the Sniffer's computer
    - Broadcast Traffic
    - Unknown Traffic (Unknown to the Switch)

  Using the SPAN protocol, we can tell the switch to mirror traffic to a particular port, based on either port number or VLAN. This means that all traffic either to/from/both will be copied and sent along the mirroring port. This allows the Sniffer to view the network traffic, even though a switch and not a hub is in use.
- *internet access*: the router has been configured to allow the users access to the UniSA network. The access will be the same as any machine within the

university, so groups can access the university's internet proxy within the controlled environment. During the exercise, this connection will remain active, but all traffic leaving will be monitored. If any traffic escapes the controlled environment, it will be blocked before or by the university's proxy server.

To analyse the network, we have a "sniffer" computer set up to monitor the network:

- *Packet-sniffing*: the sniffer is the PC running network traffic sniffer (TCPDump). It will capture all network traffic and save it in its logs for analysis. The later analysis will be using Snort and Wireshark. This PC will connect to the switch destination port, and will receive all network traffic without actually being the intended recipient.
- *TCPDump*: the purpose of this software is to capture all the network traffic on the network and save it on the hard drive for later analysis.
- *Intrusion detection*: Snort is an intrusion detection system. It can perform a *post hoc* analysis on TCPDump files which allows for a more detailed report.
- *Packet analysis*: Wireshark is a protocol analysing program, which looks at packets in some detail. This working in partnership with the other two programs allows the observer to get detailed information on some of the packets that are going across the network. This information can be analysed at a later time.

It is critical ensure there are security measures protecting the university's network from the sandpit network. On the sandpit side of the network, there are Network Address Translation translations changing internal network IP addresses to a university routable address, so the network cannot be accessed from outsid.

On the university side of the network, there are multiple security measures in place already, which we are using by plugging into their network. The port that the router connects to is protected by the university's firewall and intrusion detection systems.

There will be no additional security measures in place within the sandpit environment. This includes the sniffer PC, which will not have any security measures to prevent it being comprised during the exercise. To prevent the sniffer being compromised, there will be no IP address set on the computer, which will not permit traffic to and from the sniffer.

2.3 Staff effort

It was noted above that the peer review processed saved a great deal of academic staff time. However other staff time is used in setting up and managing groups and the network. The network in the last two years has been set up and managed by a technically-able student, as a supervised work experience or other assessed activity.

The specification of the course has been fine-tuned over the years to reduce the staff effort as well as improve delivery and correct problems. For example adding the group-level peer review represents significant saving in staff effort. In the current offering of the course, the peer review process has been streamlined with a Visitor Book where each other group must sign on at least twice a week as part of their assessment of each other group. The Visitor Book proves how often each group performed the required assessment of each other group and validates the group-level assessments.

3 The Honeypot exercise

This assessed exercise ran for the first time in the first semester of 2008, and its purpose is to build further on the skills gained during the war games exercise. There are three stages of assessment for two separate learning outcomes. The course notes including the exercise specifications are available online [4].

The first learning outcome is how to forensically examine a computing device, in our case a computer hard drive. Students working in pairs firstly had to research and decide on an appropriate forensic software tool/s, and document their decision (some starter software was also given in the notes). They were then given a "prepared" hard disc with a number of planted files and activity records was given to students, with marks allocated for their success in discovering the hidden information.

The second learning outcome is how to forensically examine a networked computer for incoming and outgoing traffic. As the students learned in the war game exercise, it can be easy for an attacker to break in to a networked computer and perform any number of damaging activities. They were grouped in groups of 5 or 6, with each group setting up what is colloquially known as a "honeypot". This is essentially a trap for hackers, which triggers an alarm if certain files are accessed. The files are deliberately planted for this purpose and should have no genuine value, but can however be made to look attractive to hackers.

This second learning outcome was assessed in three components, with a 15% component coming from the initial research on which honeypot software to set up, which was presented to the rest of the class in a seminar series. This seminar series ensured that all students benefited from the information gathered by the entire class prior to commencing the remaining work. The second component of the assessment was the group report on the group's activities, and the third component was a peer assessment worth 10%, with each member of a group assessing all other members. This is an independent group-based exercise, where the groups are not in competition with each other and do not interact in any way. One interesting feature is that the material presented in the students' seminar series was itself used in the examination.

4 Discussion

4.1 The contribution of role playing

Most of the roles played by students are common to both exercises, such as:

- *the researcher*: in many workplaces, staff are asked to research a topic for presentation to colleagues. The first component of the honeypot exercise was specified in this form, with students given a situation and role: their supervisor asked them to investigate the feasibility of setting up a honeypot so that if any intruders broke into the company's computer network, a honeypot would assist in alerting the staff as well as gathering information. The students had to investigate not just the technical feasibility but other aspects such as the legal issues and

possible social issues such as privacy, and were expected to make a recommendation on the most appropriate honeypot for the company's needs

- *the everyday user of computers*:Anecdotal feedback from students has frequently included comments about how a better understanding can help people be more aware of the need to secure their computers. It gives students a clear idea of the potential for damage or loss of private information. Students are often surprised and concerned about the ease with which unguarded computers can be hacked, as many of them are by no means practised hackers, yet in their hacker role in the war games exercise were able to implement hacking software in a short time for the exercise. The honeypot exercise also showed just how vulnerable the average computer is to hacking attempts, with students being shocked that their computers were being scanned by unknown third parties within minutes of going online. The insights gained from these exercises and the controlled exposure to attack on the real Internet has taught the students a great deal about the importance of computer security not just in the office but in the home and everywhere. Many of the students reported that during the honeypot exercise, they immediately went home to upgrade the security on their home computers.
- *the reviewer of peer effort*: staff review forms a key part of many positions in the workforce and the peer and group review components of both exercises gives the students practice in assessing the contributions of others and assessing their own contribution in light of the outputs of their colleagues. This is a serious role with credence given to their assessments. The marks awarded by students to individuals in their group and by groups to other groups are factored into the marking system. The students are expected to not just award a mark, but to explain and justify the mark awarded. The lecturer governs the final mark and awards marks independently of the student reviews for the groups' reports and each individual's report, and while trusting students to judge fairly and accurately to the best of their ability, any results that differ too greatly from the lecturer's assessment are closely investigated. Peer review does have issues, such as students being too generous or too critical of each other's efforts. They can also be quite non-discriminating such as giving everyone in the group full marks - this is generally disregarded unless the justifying comments agree with the lecturer's own observations.
- *the defender*: The most specific role is that of the defender, where students endeavour to prevent or recover from attacks from hackers. In the war games, the attackers were the other groups and hence the situation was a little artificial, but the honeypot exercise was completely genuine with unknown third parties probing the defences of the machines set up by groups.
- *team member in workplace setting*: students are assigned by the lecturer into groups, with the only requirements being a balance of undergraduate versus postgraduate students. This was to ensure reasonable skill sets amongst all groups, as Masters students often have less technical experience than undergraduates, and previous experience has found that the war games exercise made it possible for the less technically-able students to participate fully [2]. It was justified to the students as being exactly the sort of team that they might find in the workplace, with the team members selected by the company, not the team itself. Thus a major benefit of the group exercise is that it calls for a skill set well beyond that merely of technical skills. Each group requires research skills, writing skills, seminar presentation skills

and management skills, as well as willingness to work on menial tasks such as checking services of other groups. The input of non-technical people is essential to the functioning of the group, as each member of a group takes up a different role in a team, according to their skills. It places greater value on students who have lesser technical skills and encourages their participation.

Role-playing greatly improves the perceived relevance of taught courses to future employment. It is beneficial because it mimics real life situations on the job, in this case not just within employment specific to their degrees, but within any workplace or indeed everyday life. It skills students in workplace activities in any job where research is needed or where staff review is needed. It also educates students in the everyday use of their computers, whether at home or at work.

4.2 Experiential learning with these role-playing games

"Experiential learning" is motivated by a requirement for IT practitioners skilled in international collaboration [5]. Kolb has defined an experiential learning cycle [6] and this has been shown to be a valuable tool [10], and now has been mandated by the ITiCSE'99 Working Group on Professionalism [8].

The experiential learning paradigm demands that students be challenged by their learning context. Students engage in risky experiences, reflect, conceptualise, and plan. In Computer and Network Security, experiential learning is provided by the complexities of the war game and honeypot assessment exercises and the examination. Teams identifying, reflecting on, conceptualising and planning responses to the assessment challenges engage in the four stages of Kolb's [6] experiential learning cycle: concrete experience, reflective observation, abstract conceptualisation and planning.

Concrete experience: the war games exercise provides a realistically risky context for students to gain experience of both attacking and defending a network. It provides a competitive, engaging and robust sandpit in which students can develop the skills and perspective required for success in the course and later in the honeypot exercise.

The honeypot exercise provides a genuinely risky context for learning in which students observe and react to real attackers. The authenticity of the risk provides strongly convincing and often startling motivation for the course's theoretical content and its practical learning goals.

Reflective observation: providing informed critique of others' work provides opportunities to develop perspective on one's own learning and performance [9]. In Computer and Network Security, inter- and intra-team peer reviews engage students in reflective consideration of their learning experiences in order to adequately benchmark peer performance. The objectivity demanded by peer review also provides a unique opportunity for students to re-calibrate their perceptions of, and benchmark, their own performance. For many students, therefore, it counteracts the influence of ego and reduces the incidence of assessment disputes.

Abstract conceptualisation: the reflective observation phase of Kolb's experiential learning cycle [6] enables students to engage in abstract conceptualisation of their learning experiences. The course examines students on theoretical content relevant to the honeypot learning experience. Coupled with the concrete experience reported in

the seminar series, the examination provides opportunity and strong motivation for conceptualising and mastering theoretical content.

Planning: planning is a key component of two of the roles (defender and hacker) in the war games exercise. The defending role requires consideration of what attacks might be attempted and how they might be neutralised. The hacking role demands that students consider common attacks, likely defences and less common attacks. Novelty in both the defending and hacking roles leads to better peer review outcomes.

In this brief overview of how Computer and Network Security's assessment approach has actualised the experiential learning cycle, and the detailed and comprehensive learning context is evident.

We give the final word to Ted Nelson, internet personality and hypertext inventor:

> "THE GREATEST SECURITY COURSE
>
> She set up a closed computer network in a closed room, reachable only through hardwired terminals outside. Each student was given a computer on that network to defend - and from which to attack the others. Prof. (sic.) Helen Ashman's unique computer security course was the most popular course at the University of Nottingham. It elegantly combined combat with intellect and practicality." [11, p.116)

References

1. Ashman, H. (2000). Teaching Web Security Hands-on, Proc. of Ausweb 2000, Australia, SCU Press, http://ausweb.scu.edu.au/aw2k/papers/ashman/paper.html.
2. Ashman, H. (2008). "War Games" Revisited: Nine years of teaching Web security hands-on, Proceedings of Ausweb 2008, SCU Press, http://ausweb.scu.edu.au/aw08/papers/edited/ashman/paper.html.
3. Ashman, H. (2008). COMP4027 Forensic Computing, http://www.acrc.unisa.edu.au/groups/security/COMP4027.
4. Ashman, H. (1999-2007). G5CSEC Computer Security, http://www.cs.nott.ac.uk/~hla/G5CSEC (user id and password both “g5csec”).
5. Knoll, K. and Jarvenpaa, S. "Learning to Work in Distributed Global Teams." Proc. 28th Hawaiian International Conf. System Sciences (HICSS), 92-101, 1995.
6. Kolb, D.A. Experiential Learning: experience is the source of learning and development. Englewood Cliffs, NJ: Prentice-Hall, 1984.
7. Kruger, J. and Dunning, D (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments, Journal of Personality and Social Psychology 77 (6): 1121–34.
8 Little, J.C., Granger, M.J., Boyle, R., Gerhardt-Powals, J., Impagliazzo, J., Janik, C., Kubilus, N.J., Lippert, S.K., McCracken, W.M., Paliwoda, G., and Soja, P. "Integrating professionalism and workplace issues into the computing and information technology curriculum: report of the ITiCSE'99 working group on professionalism." Working Group Reports from ITiCSE on Innovation and Technology in Computer Science Education ITiCSE-WGR '99, ACM, 106-120, 1999.
9. Lundstrom, K. & Baker, W., 2009, “To give is better than to receive: The benefits of peer review to the reviewer’s own writing,” J. Second Lang. Writing, 18, 30–43.
10. Southard, S. "Experiential learning prepares students to assume professional roles." IEEE Transactions on Professional Communication, 31, 4, 157-159, 1988.
11. T. Nelson, Geeks Bearing Gifts, Lulu Press, 2008.