Paper Code

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Faculty of Engineering, Mathematics and Science

School of Computer Science & Statistics

MSc in Computer Science Hilary Term 2024 Year 1

Security of Networks and Distributed Systems
CS7NS5/CSU44032

dd MMM 2024 *Location* *Time (2 hours)*

Dr. Stephen Farrell

EXAM SOLUTION NOTES

Instructions to Candidates:

Attempt **three** questions. All questions carry equal marks (40 each).

You may not start this examination until you are instructed to do so by the invigilator.

Materials Permitted for this examination:

N/A

Exam questions are in **bold, like this.** Solution notes are in non-bold Courier New, like this.

Question 1

2024 is a year with a lot of elections happening, in Ireland and elsewhere. A commercial Irish news organisation are planning to create a new web site for election related news that will have a mixture of their own locally generated articles and articles from other news organisations. The site will allow users to comment on articles, with moderation. They plan to use a new web content management system (CMS) for the new site, showing advertising for non-paid users but with the option for a paid subscription to use the service without advertising. Paid subscriptions are a new feature for the organisation. The organisation are concerned about the risks associated with moving to a new CMS, moving from purely advertising-driven web content to also supporting paid subscriptions, (a "paywall"), and also with the potential for abuse from user comments including possibly AI-generated comments. You are the employee tasked with carrying out a risk analysis of this system, both during the planning phase and as it goes "live."

(a) Describe the process you would use to carry out this task (note: this part of the question asks about process, not the specific risks associated with the new system).

[15 marks]

The student should describe a process of identifying assets (in this case including reputation) and risks, classifying them in terms of impact and probability of occurrence, e.g. with High/Medium/Low scores for each, and assigning an overall (partial) order to the list. Normally, one then iterates, designing a mitigation for the most imporant item on the list, and then re-doing the analysis as necessary (since one mitigation may affect the probability of other risks or may introduce other risks). In practice, the process terminates when the available effort is expended. In this case, assuming good internal s/w knowledge, seeking external advice wrt risks associated with the perception of paywalls and AI-generated comments/articles would seem useful. Any answer that captures most of this is fine, as are descriptions of equally good alternative processes.

Up to 10 marks for describing a reasonable process, 5 more for overall goodness.

(b) Describe three significant risks affecting the planned system, including their potential impact and likelihood of occurrence, and outline countermeasures you would recommend that the news organisation ought apply.

[15 marks]

Up to 4 marks per risk, with 3 for overall goodness. A baseline assumption here is to require an account (paid or not) before allowing any user comments, sugesting otherwise is a negative. Risks here are fairly obvious, but examples include:

- The new CMS may be buggy: check for history of CVEs and their handling, have a site-status X/insta/FB page, etc
- If the new CMS has plug-ins, those may be buggy or badly maintained; choose carefully! And keep an eye on developments. (Think wordpress, but don't do that:-)
- Admins might misconfigure the new CMS; get training
- User comment moderation might be expensive and/or error-prone (either over- or under-blocking); pay moderators properly, anyalyse their actions, setup policies ahead or time that are user-visible and maybe even selected with end-user input
- Those involved in elections may try dominate or spam the comments on particular articles; have an AUP disallowing that, call it out when it happens (that last mitigation may add controversy)
- Partner organisations might provide AI-generated articles; make it a contractual requirement that they do not.
- Users might dislike the idea of a paywall, reducing site-visits and hence advertising revenue; explain why the paywall makes sense for small news orgs without scaring the punters too much about the real horrors of advertising (RTB etc.)
- Non-paid users may use adblockers; provide some real added value for paid-users.
- Competitors or overly-enthusiastic political fans might DDoS the site; use a good CDN for enough of the site to make that hard to have significant impact.
- Outsourcing, via a CDN, or a moderation provider, risks various attacks; contracts and picking good providers that are audited

(c) What risks that potentially affect end users should be considered, and what actions would you recommend be taken to mitigate those risks?

[10 marks]

5 marks/risk, assuming one or two provided, e.g.:

- Phishing attempts based on user account maintenance emails; don’t send such mails & make it clear to users they won’t get any real ones
- Non-paid users without adblockers are likely exposed to the RTB system. Probably no mitigation from our news site.
- Credential stuffing attacks here would be attractive to political parties or their operatives/supporters given an account (paid or not) is required for commenting. Provide some 2FA option, maybe a “Login via gmail etc.” and password strength indicators
- Unusual/inventive risks get a bonus, if realistic (here and everywhere)

Question 2

(a) For any real Internet key exchange protocol (e.g. TLS, IPsec, or Kerberos) describe the key management and application data protection aspects of the protocol in detail. Describe how the cryptographic mechanisms used by your chosen protocol may affect server performance at scale.

Your answer may describe any widely used version of your chosen protocol, but for TLS, you must explicitly state which version you are describing.

[25 marks]

Marks-from-15 for a good description of the scheme, with e.g. the TLS handshake and application layer properly described. 5 marks for describing possible perf issues (e.g. 25519 vs. p256 vs DH; use of AES NI etc.) and 5 for additional goodness (e.g. covering optional TLS extensions/tickets etc).

Historically, 95+% of students will very reasonably choose TLS and those that do not tend to be well clued-in about whatever else they pick (usually IPsec or Kerberos 'cause they've done some work on that)

(b) Describe an application that can use your chosen protocol and that would be at risk in the face of a cryptographically relevant quantum computer? What application-layer mitigations and/or changes to your chosen protocol would you suggest to handle those risks in the short term (this year), and in the medium term, e.g., in five years' time?

[10 marks]

5 marks for describing an application that suffers today's PQ problem properly (record-now-decrypt-later) and other PQ issues (e.g. immaturity of algs, IPR on Kyber, ...); Example could be anything processing genomic or health data. 5 marks for covering hybrid TLS and justifying short/medium term answers (which can vary depending on application).

Note – this part of the question is exactly the same as last year – but that's justified as the topic is as or more relevant to understand this year, which trumps being unpredictable in this case.

(c) What kinds of issues might arise should a server using your chosen protocol be deployed on a small device, such as a home controller running on a home network?

[5 marks]

Straight marks from 5 for overall goodness, issues, assuming chosen protocol is TLS, may include:

- the BIG one: no easy/generic way to get a TLS server cert for 10.0.0.x or 192.168.x.y
- if one swaps out either client or server you likely need to reconfigure the other one from scratch
- UIs for such devices tend to make complex configuration very hard
- if client and server from different vendors, ways to do (re-)configuration may not be commensurate
- client might assume there’s a working DNS name for server, bringing mDNS dependencies into the picture

Question 3

A regional transport authority (such as Transport for Ireland) provide a mobile application that allows users to see bus/train timetables and real-time information, and that is very widely used in the region. The application currently has a minimally invasive privacy policy, providing the option to access information without an account or without requests being easily identified as being from a particular device. Users do have the option to create an account, e.g. to get email notifications, but few users enable this option.

The local government and the transport authority now wish to enhance the mobile application to help them better plan changes to transport routes and to monitor pedestrian use of busier city-centre streets. This may require them to do some form of (possibly per-user) tracking, so they can accumulate information about, e.g., peaks and troughs in pedestrian use of streets, the numbers of people making specific bus/train transfers etc. However, adding such functionality clearly creates new potential privacy issues and may not be well-accepted by users.

You are tasked with designing a privacy-friendly enhancement to the mobile application and related back-end systems to achieve the goals above.

(a) Outline a design for the above enhancement (include diagrams as appropriate), and state the security and privacy requirements the resulting overall system should meet. Note that this part of your answer should only discuss security and privacy requirements and not describe specific mechanisms to meet those requirements.

[20 marks]

8 marks for overall design, e.g. privacy-friendly potential; 8 marks for more obvious requirements identified 4 more for overall goodness. Basically some obvious design options exist here, from full location-enabled snooping on a per-user basis, to inferring approx statistics from existing traffic (if asking about Dublin bus/stop 400, user is likely near TCD/Pearse St.), to complicated privacy-preserving metrics (not covered in class, so not expected here). Privacy requirements relatively obvious (less invasive == better), for security, provider doesn’t want users to be able to spoof one another, esp not in (fake) crowds, otherwise usual mobile web app + servers stuff (e.g. DDoS resilient).

As usual some students will mix up requirements & implementation/mechanisms (i.e. parts (a) & (b)) but I’ll be lenient on that.

(b) Describe, in detail, the most important security and privacy mechanisms you would propose for the system to meet the requirements from part (a).

[15 marks]

Marks from 10 for workable and privacy-friendly answers, or, for good justification of less privacy-friendly answers, with 5 more for inclusion of less obvious things. The privacy spectrum here is roughly, from:

- ask to turn on accounts & GPS and track every app user as much as possible reporting full detail to servers while downplaying the invasiveness of that to users in public statements and policy text

  to:

- arrange someone independent and "trusted" (e.g. an EFF-like body such as ICCL here in Ireland, privacy international in the UK maybe), to act as a proxy for receiving privacy-preserving metrics as to which of a few named streets/bus-stops one was "at" sometime in the last N hours – then have the back-end servers make queries on that proxy for statistics.

There're lots of possible designs here, basic marking idea is more marks the more practical and privacy-friendly the mechanisms, *or* the better the justification for use of privacy un-friendly mechanisms.

(c) What are the best arguments you can provide for a) deploying the enhanced mobile application as described, and b) for not deploying an enhanced mobile application at all?

[5 marks]

It's ok if the arguments provided are at the political/policy level and not at the technical level. Most marks if both pro- and anti- arguments are equally good.

There's an interesting argument, that, if made anywhere, will garner some bonus points: the more a mechanism is really privacy-friendly the harder it probably is to explain to users, (e.g. k-anonymity, various crypto things), making it, very quickly, harder for users to distinguish services really trying to do the right things for privacy, versus those services that only emit the canonical "we care deeply about…" while actually doing the opposite.

Question 4

(a) Describe in detail how DNS query resolution for typical address records works, including, at each stage, potential attacks and mitigations. How might the new HTTPS resource record type affect browser performance if/when that is very widely deployed.

[15 marks]

Marks-from-7 for a reasonable description of the stub, recursive, authoritative resolver actions with/without a cold-cache. 5 marks for documenting attacks/mitigations for each stage with e.g. a good description of DNS poisoning getting 3. 3 marks for knowing how HTTPS RRs might affect performance, maybe a bonus if they note the A RR vs. HTTPS RR ipv4hint duplication problem for stubs that run a race between queries for A/AAAA and HTTPS. (IOW, seeing the problems we try solve via happy eyeballs.)

(b) Describe the trust model and operation of DNSSEC, and the typical steps involved in validating DNSSEC-signed DNS responses. Explain why the number of domains for which DNSSEC is available is so small.

[15 marks]

8 marks for describing DNSSEC well and 2 more for extra goodness, e.g. if they included NSEC or CDS/CDNSKEY or similar. 5 marks for reasonable description of why we have single-digit DNSSEC deployment – the DS at parent issue, registrant→registry comms required, KSK rotation etc.

(c) How do privacy issues arise in connection with the use of the DNS? How are those affected by the use of DNS-over-TLS (DoT) or DNS-over-HTTPS (DoH)? What default DNS privacy setup would you recommend for a current mobile operating system? Say why you consider that a reasonable default, and how it would affect those privacy issues.

[10 marks]

4 marks for describing privacy issues well (act of accessing name from IP at time, sets of qnames being identifying etc.). 3 marks for correctly stating the effects of DoT/DoH (with maybe a bonus if done really well, e.g. noting that there are displacement issues at play) and 3 marks for a well justified non-crazy default, which may well be Do53 'cause DoT as an OS default does imply some failure cases we might not yet be willing to bear.