P5

Interviewer: Let's get started. You can share an example of a code auditing task you've done before and explain how ChatGPT helped in the process.

P5: In a very traditional audit, the thinking is all pretty much the same. First, look at the general whitepaper or some relevant materials about its technical architecture to facilitate a general understanding of what the project is about, right? Then read through the code, look at the framework and so on, and then look at those key functions and such. First, just do a general review of the entire framework to figure out what the project is mainly about and which aspects are relatively important for this project, such as those involving funds or some upgrades and the like. Pay attention and mark them. Then take a closer look. After getting an impression from the first reading, the second reading focuses on some important function interfaces, right? And then the entire process of some variables. At the beginning, some variables are defined, and then check if there are any possible abnormal situations in the state changes during the execution of the entire set of functions.

Right, and then when conducting an audit, after marking those items, we should first list the high-risk issues that can be identified, and finally write the report. The reason for this report is that it not only includes high-risk issues, nor does every project have high-risk issues. Generally, there are relatively more medium-risk and low-risk issues to address. Based on some experience and common sense, we need to do some sorting of info and the like. If we use fewer tools, in fact, I use them relatively less. Because for static analysis now, it may be helpful for that kind of info, but there are fewer simple ones, and I often don't use it because that kind of info can be seen at a glance.

Interviewer: Is that information related to information?

P5: Right, it's mainly about optimizing some cases or providing suggestions on code style. If we use ChatGPT later, some prompts can assist in doing this. On the one hand, it can output audit results, and on the other hand, it can provide some interesting points that we might not think of, which may open up some new ideas. However, its false positive rate is still quite high, but some of them can inspire ideas and provide assistance. Additionally, it can output in a certain format. For example, if you describe something to it, it can output the relevant description or format of the code, and you can directly use it as an assistant to paste some code into the report. Generally, it mainly helps with output format and content.

Interviewer: Assistance with output format and content. Does it refer to the format of the final report?

P5: Right. You can give it a template, then paste the code and describe it, and then it will give you a specific template and specific output based on the template.

Interviewer: Just now you mentioned that you look for vulnerabilities, which means these vulnerabilities come in different types, right? Then these different types of vulnerabilities are, so to speak, some you assign to GPT to look for, while you look for others yourself. Let me rephrase my question. You just mentioned different types of issues, such as some being variables and some being function functionalities. So which ones do you think you can rely on GPT to help you find, and which ones do you prefer to find on your own?

P5: For this, I haven't fully studied and understood how to use ChatGPT to handle these things. I can only give it a good prompt, then provide it with the code and ask it to output some content. But when it comes to having it do specific tasks like sorting out the state changes of its variables, I may not have attempted that yet.

Interviewer: So when you use ChatGPT now, it's just equivalent to a single-turn conversation, right? For example, you input code to it, then let it directly give you an output, and then you'll ask step by step like this to guide it to find some bugs.

P5: Right.

Interviewer: Are you referring to the second type, or the single-round dialogue?

P5: I'm the kind of person who, when given the code, either can switch it or, if you're not satisfied with the result, can retry. I'll just keep retrying to see if I can get something effective.

Interviewer: You're not saying that, for example, after you give it something, it generates a result, and then after you look at this result, you continue to ask further in-depth questions based on this result, right?

P5: Some will. If you can tell at a glance that it's a false alarm, there's no need to let it retry; just let it provide a different answer and don't dig deeper. If it's ambiguous but somewhat relevant, you may need to ask. Right, but basically, I only make a judgment when it's relatively certain. If it's ambiguous, I'll ask a bit. If the answer is still rather vague after asking, I'll directly retry and let it take a different path.

Interviewer: Have you tried using ChatGPT at different stages, for example, when we divide code auditing into different stages?

P5: It might be the two aspects just mentioned, namely the final result output and then the organization of some of his issues in the middle. Regarding what you said earlier about using it for sorting out some basic information of projects, I may not have done that.

Interviewer: You haven't used ChatGPT for information organization. Why is that?

P5: Maybe didn't expect it, didn't think of doing it this way before.

Interviewer: If you were to use it now, you can imagine how you would use it. You can fully imagine, and then how you can make full use of GPT during this process.

P5: Just throw the white paper of that project at it, let it study it first, and then ask questions using code. This might be a bit better. Previously, we just gave it the code directly, which might have been more difficult to read.

Interviewer: What was the purpose of having it read the white paper?

P5: Let it have some understanding of this project, because that's how my current audit process works, so it and I...

Interviewer: So your goal is to have it simulate that process. Because your process is equivalent to first having a more high-level understanding, and then looking at some details, right? You hope GPT will do the same. What kind of goal do you hope to achieve through this?

P5: If I were to say something about the auditing project itself, reading the white paper might only give me a limited understanding of the project. In fact, I may not have a very in-depth understanding of it. For many of those relatively professional aspects, they may only be briefly mentioned. If I were to let GPT analyze it, I think its understanding might be deeper. I feel this way because I haven't tried this before.

Interviewer: What specific professional knowledge are you referring to?

P5: For example, some mathematical models, economics, etc. This may be because they involve some knowledge related to mathematics, and if you want to calculate, it will be very time-consuming.

Interviewer: Do you think you've learned something during the process of using it to help you understand? Have you learned something new, or is it just a supplementary tool and you already had this knowledge?

P5: I probably didn't learn much because it just gave me a result directly, and I have no way of knowing his specific thinking process. It would be better if it could output how his solution came about.

Interviewer: So you hope that it not only gives you a result, but also provides you with some explanations.

P5: Right.

Interviewer: But will you read these explanations carefully once they are provided?

P5: If his accuracy rate is relatively high, it is still worth learning. However, at present, his learning cost is relatively high, and his false alarm rate is also relatively high. After learning, you may not even know whether what you've learned is correct, which is a bit troublesome.

Interviewer: So you don't quite trust some of his results, right?

P5: Right, the main decision still has to be made by oneself. Because sometimes what it says sounds very reasonable, but upon reflection, there are still significant problems.

Interviewer: So you use some other software, such as VSCode extensions, or some online (audio) auditing platforms.

P5: It's auxiliary auditing, right? It's used relatively infrequently.

Interviewer: What's the reason?

P5: The reason might be that I haven't had much exposure to tools. Audit tools are relevant, or auxiliary relevant, and it could also be that there aren't currently any good ones, or ones that I'm used to using.

Interviewer: Okay. So you think using ChatGPT actually doesn't help you save any energy or time, right? And how do you think its performance compares to using traditional methods?

P5: Sometimes, ChatGPT can be helpful, for example, when you don't have much time or the audit period is relatively short. It can directly provide you with some answers, and you just need to directly judge whether they are correct or not. In this regard, if the audit schedule is tight, it still has a certain effect.

Interviewer: So you said that you would use it when you're in a hurry.

P5: Also, you've basically completed the audit, and then see if it can provide some very different perspectives of thinking.

Interviewer: I find this part quite interesting. Could you elaborate on how you used it to gain multi-angle thinking?

P5: If you've already reviewed this project, for example, the reports are basically completed, and there's still some reserved time left, then you feed that information to it, and let it output different justifications. Then you take a look. Sometimes, although it may make mistakes, the scope of its thinking might broaden your perspective. Right? If you follow its imagination, you might make different discoveries, right?

Interviewer: So after you finally generate the report, you not only want it to check whether the report is correct or see if it can help you improve it, but more importantly, you will also ask it if it has any other ideas, right? That's how you use it.

P5: If the actual construction period is relatively tight, then it helps me output some things.

Interviewer: When it's relatively urgent, the output provided to you should have false alarms, right? At this time, doesn't it actually reduce your accuracy?

P5: Right, those relatively obvious false alarms can be directly investigated and eliminated, and then for those that are clearly confirmed at first glance, you can briefly write them down.

Interviewer: So it can still be of some help if you use it.

P5: It is relatively labor-intensive to investigate false positives.

Interviewer: When you use ChatGPT to assist you in code auditing, what do you think is the biggest challenge you've encountered? It doesn't have to be that serious; just some difficulties you've faced or areas you think are not very user-friendly, aside from the false positives just mentioned.

P5: Let me think. Also, sometimes, for example, after you give it a piece of code, and it has many questions. After you ask question a, when you then ask question b, there is a process equivalent to an interruption at that point. Then if you ask question b, it will combine the answer to question a to give you the answer to question b. This might be rather troublesome, and you need to take a new table to ask question b.

Interviewer: So it's pretty much like this, right? For example, after having several rounds of conversation with it, you suddenly want to ask another question about the previous one, but also want it to have the context, yet it thinks you want it to continue from the previous question.

P5: Right.

Interviewer: I think this is quite interesting.

P5: In this case, we have to create a new table to ask it.

Interviewer: But if you start a new conversation to ask it, you'll have to train it again.

P5: Right, so this is rather troublesome. It might also be that the way I asked led to his misunderstanding. I'm not quite sure either. Anyway, it's always difficult to handle whenever this happens.

Interviewer: Understood. Are there any other areas you find inconvenient, not only when using ChatGPT but also when using traditional methods?

P5: They're all pretty much the same. They share a commonality in addressing the issue of false positives. Whether it's traditional tools, which also have many false positives, or GPT, both will have them.

Interviewer: Is this a major consideration for you not to use these tools?

P5: On the one hand, and on the other hand, the things it reports are not very good. Either they are relatively basic things that you can tell at a glance. If the logicality is relatively high, it fails to recognize them, which is very contradictory. That is, if you use it, it feels inferior to doing it yourself, and if you don't use it, you haven't really used it.

Interviewer: Speaking of this, I'd like to ask if Mr. Xue is here?

P5: Mr. Xue?

Interviewer: Regarding the point it just mentioned, I'm not sure what solutions you, as an expert, would have from your perspective. How should I put it? ChatGPT fails to detect some of the more complex vulnerabilities it wants, and for the relatively simple and basic ones that ChatGPT does detect, it can actually spot them quickly himself. So I'm not sure if this is a problem specific to junior auditors or if experts also face it.

P5: I think this is actually a problem with ChatGPT, and any use of ChatGPT may encounter this issue. As for how to solve it, currently, it is usually done by optimizing the prompt. Generally speaking, I have roughly gone through the following stages. The first stage is pure questioning. If it's just pure questioning, the problem is quite significant. When the problem is significant, it will actually lead to the issues mentioned earlier, such as hallucinations or the inability to effectively output real vulnerabilities. This is the first stage.

In the second stage, I optimized the prompt. Actually, this is the same prompt I gave you before, to someone I don't know. That prompt is much better, but it still has some issues. Let's first talk about where it's good. That is, its output is no longer just certain non-vulnerability-related things; it has become more effective. Another point is that it can conduct some relatively in-depth thinking based on what the vulnerability reflects. However, there are some drawbacks. To put it simply, this drawback is that it still sometimes outputs some ineffective things, which is unavoidable.

The second point is that it often makes some background knowledge errors. For example, transactions on the blockchain are principled, which means they either succeed or fail. But it's obvious that GPT often forgets this when analyzing transactions. It thinks that transactions can be interrupted, can be interrupted by multiple executions, and can also be interrupted by multiple transactions. This is a very serious problem, which is equivalent to me having to popularize some background knowledge for it. So this is also why I added background knowledge to it in the later stage of the prompt. Another shortcoming is the first problem mentioned earlier, which is that sometimes what it outputs is not really a real vulnerability. This is one aspect.

Another aspect is the third cycle, which means that what I am currently doing is that I no longer rely on unknown things, nor do I rely on the capabilities of GPT itself. What I rely on is my Knowledge Base. In comparison, it is more effective. At this point, it will no longer output those meaningless loopholes. However, in fact, we still encounter some problems now, and there will be more problems in the future. We will still encounter hallucinations, knowledge gaps, and various issues.

Interviewer: We plan to develop a tool to assist auditors in better interacting with ChatGPT, and through this improved interaction, help auditors better identify those vulnerabilities. So, do you have any ideas about this tool or expectations for its functionality?

P5: I think this is really good, equivalent to creating a customized GPT specifically for auditors. Actually, considering from my own experience, regarding the coherence we just mentioned, it's like you can input an entire project, and it will automatically output everything, whether it's the project background, the overall execution process, or the specific audit results, right? There will be a general overview, and then you can ask specific questions, which is quite good.

Interviewer: First, let it give you a general framework understanding, and then proceed to local understanding.

P5: But actually, in my understanding, this process is more like a kind of Modularization analysis, or for example, although a project is very large and cannot be analyzed all at once, it can be split into various small files, and then finally we combine the conclusions of these small files, which I think might be achievable.

Interviewer: Is it something like generating a logic diagram, which contains different modules and indicates the relationships between them, and then letting GPT do something?

P5: Right, after expanding, then for some minor points, we can ask GPT about different points. This is still quite helpful for auditors. Moreover, I think there's a rather crucial aspect here, which is the issue of interactivity, that is, it can resolve a previous concern of mine. If questions about a and b are asked in a cross manner, it may not be very clear. With this interactive approach, for example, if you ask in detail about each function of contract a and each function of contract b, there won't be any cross - referencing, and it won't mix up the answers when reading them.

Interviewer: Indeed.

P5: Actually, there's another form that I think is also quite good, which is a VSCode plugin. It's like adding comments for you and so on.

Interviewer: Provide explanations, right? Provide explanations for different Code Blocks.

P5: Right, it's like adding comments. You can ask questions based on the comments, and this kind of approach might be more convenient.

Interviewer: What kind of tool, specifically what kind of effect it achieves, would make you really want to use it?

P5: In its final form, it will directly generate a report for you. For the intermediate form, it is capable of, for example, answering detailed questions. After expanding this, it can accurately describe all details, and when asked questions related to vulnerabilities, it can also provide relatively accurate descriptions.

Interviewer: So you think accuracy is one of your priorities.

P5: Right, as long as the false alarm rate is not too high, it's okay. But false alarms are definitely allowed. However, if the false alarm rate is relatively high, it may still take too much time to address those false alarms.

Interviewer: I'd like to ask, for the auditing work, do you do it on a per-occurrence basis, a per-project basis, or is it something that's continuously ongoing, like how development is continuously carried out?

P5: Right, I should be, let me think, right, it should be engaging in your position. Since you're an auditor, there must be audit tasks, and of course, there are also some other things, some other related ones.

Interviewer: You will also have other jobs.

P5: Mainly still doing auditing.

Interviewer: I think the discussion just now was really interesting. Thank you.