HR001119S0076
Securing Information for Encrypted Verification and Evaluation (SIEVE)
Frequently Asked Questions

As of July 23, 2019

Q18: The BAA specifies that in Phase 3, TA1 approaches that address proving statements about programs will need to reason about programs with statements that are "Probabilistic Events." We could imagine a few interpretations of this goal; namely, programs may perform actions that draw from some specified distribution. A TA1 approach will then need to communicate proofs of properties of at least one of the following forms:

1. Guarantees about the distribution of possible program outcomes (e.g., the program is secure except for a negligible number of cases);

2. Guarantees about every possible program outcome, basically treating values obtained from probabilistic events as chosen non-deterministically.

Which classes of program properties are in scope for the program?

A18: Either would be in scope for TA1.

Q17: The BAA states that TA2 should provide a collection of ZK proof mechanisms that support various trade-offs between, e.g., prover complexity, verifier complexity, and overall complexity. However, the BAA also specifies that in Phase 2, TA2 approaches need to achieve an asymptotic overall complexity of O(n + k), and likewise, a communication complexity of O(sqrt(n) + k) in Phase 3. How important are these asymptotic metrics versus providing (practically efficient) ZK proof mechanisms that explore the full complexity tradespace even if they do not necessarily achieve the required asymptotic metrics?

A17: TA2 performers should state whether their solutions will be able to meet the metrics as set out in the BAA. It is anticipated that, within the collection of ZK proof mechanisms provided, some mechanisms may require tradeoffs that are less efficient in some parameters (e.g., prover computation, verification computation, communication). The metrics were selected to ensure that TA2 performers can handle very large proof sizes efficiently *at a minimum*; additional ZK mechanisms proposed do not have to meet the metrics but must justify why such mechanisms would be optimal in some DoD-relevant scenarios.

As of July 17, 2019

Q16: In social settings, we may have multiple players with different motivations, not just “honest/adversarial”. Is that within scope for TA1?

A16: Honest/adversarial appears to be more of a security model question, which would be more in scope for TA2.

Q15: Translating legal notions, such as GDPR compliance, or social notions, such as "accountability of the court system even when records are sealed," requires collaboration with legal experts. How much is such translation in scope for TA1?

A15: This is completely in scope for TA1

Q14: There have, in the past, been intermediate representations that are somewhat specialized, like Camenisch-Kiayias-Yang statements about properties of discrete logs, that allow for very expressive, but not all of NP, statements. Will such IRs be considered? Or is this mostly Boolean/arithmetic circuits?

A14: Such representations are in scope; however, TA1 proposals should motivate the usability of their encodings by TA2 performers.

Q13: Can you elaborate (or re-state) the idea of ZK proofs about probabilistic computations?

A13: Some DoD-relevant problem statements may encapsulate probabilistic events (e.g., heap spray attacks, random walks, non-deterministic computations)

Q12: How important are "security proofs of the provided systems" for TA2? Are these computer-checked proofs or on paper.

A12: TA2 performers should specify how the security of their solutions can be verified (e.g., via human- or machine-readable proof, etc).

Q11: Can there be a basic IR standard, e.g., Boolean circuits, and extensions thereof that TA1 teams may or may not decide to use.

A11: The standard is intended so that any TA2 performer can, in principle, implement any TA1 output. TA1 proposals should highlight if they do not believe that TA2 performers will be able to implement their encoded statements into a zero knowledge proof.

Q10: What is the timeline for converging towards the IR standards?

A10: Per the BAA, a draft standard will be in place by the end of phase 1.

Q9: How will the TA2 performers "lead" the IR standardization effort given that there will be several TA2 awards.

A9: This process will occur collaboratively.

Q8: Can you explain to what extent TA2 proposals need to encompass any possible IR?

A8: TA2 proposals may discuss which IRs they plan to handle, and the tradeoffs (e.g., efficiency, coverage of different classes of problem statements) associated with their approach. Technically, the ability to create a ZK proof with respect to an NP complete IR implies the ability to handle any other IR in NP.

Q7: Are offline computations and complexity included in the metrics?

A7: Yes, though amortization is something that can be discussed in the proposal.

Q6: To what extent does the generation of IR encodings need to be fully automatic or can it have manual steps

A6: It can have manual steps (though automated seems more broadly applicable)

Q5: Are slides going to be available?

A5: Yes back on the BAA announcement site (FBO)

Q4: What is the total program budget? What is ballpark budget per TA?

A4: DARPA is not releasing this information at this time. Proposal cost should be proportional to level and scope of effort.

Q3: Can you give an example of what you mean by "Interplay of problem statement complexity vs security leakage"?

A3: Some problem statements may reveal more information about the underlying sensitive information than other statements. For example, a statement about the existence of a camera on Street A in City B would leak more than a statement about the existence of a camera in City B.

Q2: Is a graph a valid IR for TA1? Or do you want it compilable (e.g. bytecodes, llvm, circuits)?

A2: TA1 proposals should motivate how their outputs will be usable by TA2 performers.

Q1: Can you elaborate on the IR standardization discussed? To what degree can TA1 performers develop an IR optimized for their particular domain vs mandatory adherence to program wide standards?

A1: TA1-encoded IR outputs should be able to be processed by any TA2 performer, although some TA2 zero knowledge regimes may be more efficient than others. The standardization process is intended so that all TA2 performers can more easily process on any TA1 generated IR encoding of a problem statement. For example, all Boolean circuit encodings should be in the same format across all TA1 performers.