Verifying Payment Channels with TLA$^+$

Matthias Grundmann

Bitcoin introduced a way to perform decentralized payments without a trusted third party. However, Bitcoin does not scale in the number of transactions performed per second. To improve on this, Payment Channel Networks have been proposed that allow parties to securely perform transactions off the blockchain: Two parties create a payment channel by locking funds in a shared account on the blockchain. Both parties keep the state of how the funds inside the channel are distributed between themselves. The two parties can perform transactions off-chain by updating the ownership of the channel's funds in their shared state. Each party can close the channel at any time by publishing the latest state on the blockchain which will make the final balance available.

The security model of payment channels assumes the counterparty to be untrusted and adversarial. A dishonest party might close the channel with a favorable outdated state. A protocol for payment channels needs to ensure that such dishonest behavior can be punished. More general, a payment channel protocol should guarantee that a party will finally receive their correct balance on the blockchain if it follows the protocol. Developing a protocol for payment channels such as the Lightning Network [1] is a challenging task because the proposed protocols are complex and many edge cases need to be considered. Our hypothesis is that we can show aforementioned guarantee for the Lightning Network's protocol by model-checking a TLA$^+$-model of the protocol.

In our project, we model the Lightning Network's protocol in TLA$^+$. The challenges we faced in this project include: (a) Building a model of the underlying blockchain [2] and transactions that models hashes and signatures. (b) Modeling time and allowing a dishonest party to deviate from the protocol while still keeping the state space explorable. (c) Providing a protocol developer with an intuitive and understandable output in case a counterexample is found. To approach these challenges, (a) we developed a model of transactions that can be reused for models of other off-chain protocols, (b) we modeled the adversary in a way that adversarial actions that are trivially invalid are excluded from the model, and (c) we created a visualization of state traces so that flip-book style animations can be created visualizing executions that break the security guarantees.

With our project, we aim to build a method that can create confidence in the correctness and security of off-chain protocols and that can help developers to quickly find flaws in proposed protocols. With our talk, we hope to draw interest of the TLA$^+$ community to this decentralized payment use case, to gather feedback on how to improve our model, and to discuss how to address open challenges.

[1] Joseph Poon and Thaddeus Dryja. The Bitcoin Lightning Network: Scalable Off-Chain Instant Payments. Technical report, 2016.

[2] Matthias Grundmann and Hannes Hartenstein. Fundamental Properties of the Layer Below a Payment Channel Network. In *Data Privacy Management, Cryptocurrencies and Blockchain Technology*, Lecture Notes in Computer Science, pages 409–420, 2020.