An epistemic view of security properties

Musard Balliu Roberto Guanciale Matvey Soloviev

KTH

{musard,robertog,matvey}@kth.se

Various security properties, from basic ones such as confidentiality and integrity to more involved ones such as robustness and nonmalleability, have been defined and studied in the security community. Generally, the definitions of these properties are made with respect to particular system models and formalisms, which are often tailor-made to the problem at hand and therefore make it difficult to compare and understand the differences between and implications of various definitions. Epistemic logic is a type of modal logic that is used to model the knowledge of one or more agents which are uncertain about the state of the world they are situated in. It stands to reason that such a logic can be used to reason about confidentiality, that is, the property that secrets do not become known to agents who are not authorised to know them. By augmenting a standard epistemic logic with modalities representing the capabilities of attackers and other agents in a computer system, we show that we can represent a variety of security properties in a way that is intuitive and agnostic to the details of the particular setting studied. Among others, this enables us to deepen our understanding of the properties, validate enforcement mechanisms in unusual settings and connect them to existing work on reasoning under uncertainty.

We represent a computer system as a multimodal Kripke frame $(\mathcal{W}, T, K_A, R_A, W_A, \ldots)$, where the set $\mathcal{W}$ of possible worlds contains all possible states that the system could be in at a given point in time. This enables us to naturally reason about how knowledge and effects develop over time. Previous approaches we are aware of, such as the one of Moore, Askarov and Chong, generally took a single world to encode an entire possible run of the system. On top of this, we define

- The "time" relation T, which relates two points w and w' iff the latter can be obtained by letting the system in the former run for some number of steps;
- The "allowed knowledge" relation K_A for each agent A, which relates two worlds if A can't distinguish them based on what A is allowed to read;
- A relation W_A that can be seen as an "allowed write" relation, relating two worlds if they only differ in memory that A *is* allowed to write;
- The "program change" relation R_A, which relates two worlds if they only differ by A's choice of program.

In this setting, we can then for example define confidentiality as the property that at all worlds w, for all agents A, for all simple Boolean formulae φ (representing a fact that can be known),

$$w \vDash \langle T \rangle [K_A] \varphi \Rightarrow [K_A] \langle T \rangle \varphi,$$

that is, if A will eventually know that φ is true, then A must already know that φ will eventually be true. In other words, A must not come to be able to infer any new knowledge from the information that A has access to.