Probabilistic Verification of BGP Convergence

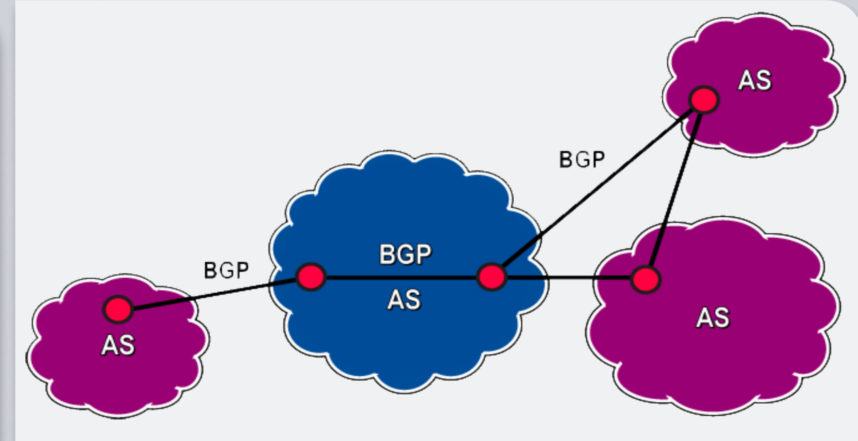
Soroush Haeri
Simon Fraser University
Vancouver, British Columbia, Canada

Dario Krešić
University of Zagreb
Varaždin, Croatia

Ljiljana Trajković
Simon Fraser University
Vancouver, British Columbia, Canada

BGP Convergence

- Border Gateway Protocol (BGP) is widely used as the main inter Autonomous System (AS) Internet routing protocol.
- An AS selects its preferred routes based on its routing policy and the best routes that have been advertised by its neighboring ASes.
- Local AS policies play an important role in preferred route selection because the BGP allows policy-based decisions to override distance metrics.
- Local routing policies are usually defined based on a limited knowledge of other AS policies and network topology and, hence, may be inconsistent.

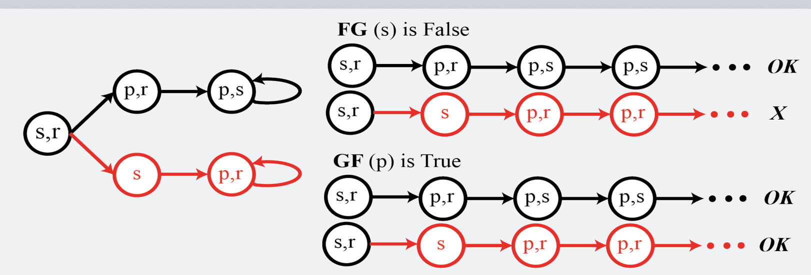


Routing among BGP systems

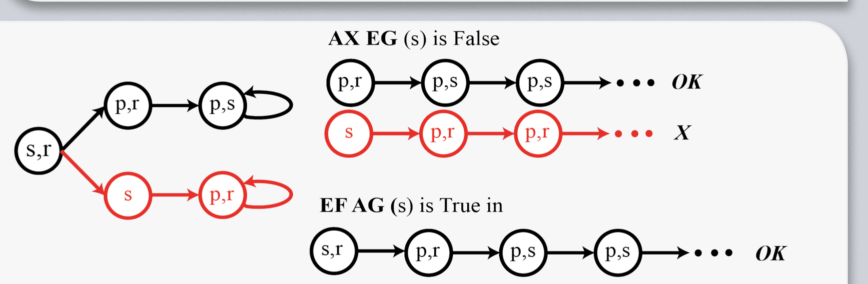
- These policies may cause a set of ASes to exchange route information messages indefinitely and not to converge to a set of stable routes.

Model Checking

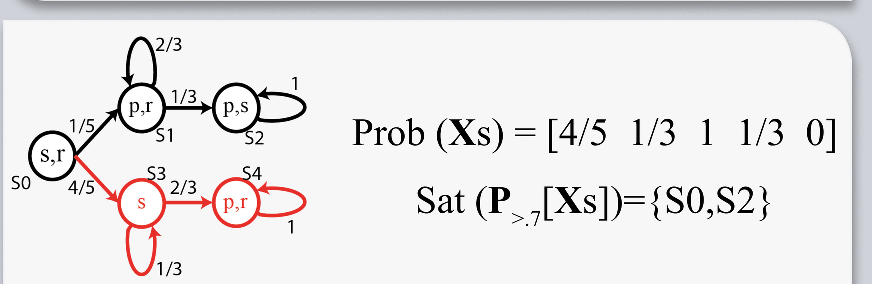
- Model checking is an automated technique to formally verify the correctness of a finite-state system.
- Safety and liveness are two important specifications for communication protocols.
- Liveness is a desired property that should eventually happen.
- Input to model checking process is a variant of finite-state systems and required specifications are expressed in terms of temporal logic.
- Linear Time Temporal Logic (LTL) encodes information about the future of paths.
- Model of time in Computational Tree Logic (CTL) is a tree-like structure.
- Probabilistic Computational Tree Logic (PCTL) introduces probability operator to CTL.



LTL model checking



CTL model checking



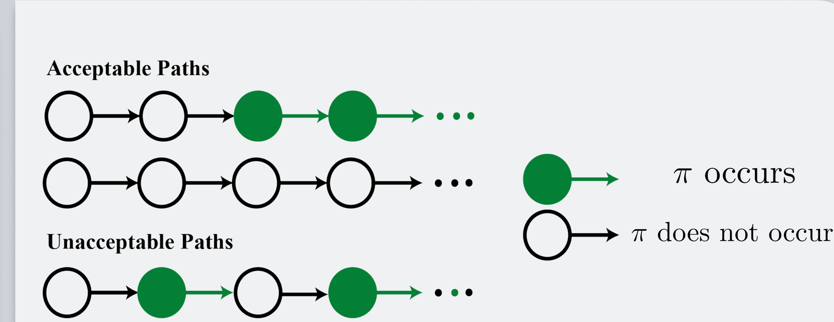
PCTL model checking

Global BGP Execution Model

- Viswanathan et al., describe the global BGP execution model as an input-output automaton.
- Assumption: node 0 is the single destination for all other nodes.
- Assumption: $Q(S)$ is the set of states of the global automaton describing path assignments.
- Let π be a mapping function that assigns each node u to a permitted path $\pi(u)$.
- For every node u, the path assignment initially maps the empty path ϵ to u $(\pi(u) = \epsilon)$.
- Inputs to the automaton are events of the form: $\{advertise_u \mid u \in U\}$, for some $U \subseteq V - \{0\}$, where V is the set of all nodes.
- The transition matrix $\mathbf{T}(S)$ of such automaton is of dimensions $|Q(S)| \times |Q(S)|$ and $\mathbf{T}(S)_{ij} = \{U \mid \pi_i \overset{\{advertise_u | u \in U\}}{\longrightarrow} \pi_j\}$.
- Let $\mathbf{p} = (p_1, ..., p_n)$ be an activation probability vector with $n = |V| - 1$, where each p_i represents the probability that node i receives an event $advertise_i$. Node i recomputes its routes after receiving the event.
- $\mathbf{T}(S)$ may evolve to a stochastic transition matrix $\mathbf{T}'(S)$ by casting operator $P(\cdot)$ on every element of $\mathbf{T}(S)$. Let γ be a subset of power set of $V - \{0\}$. $P(\cdot)$ is defined as:
$$P(U) = (\prod_{i \in U} p_i) \prod_{j \notin U} (1 - p_j),$$
$$P(\gamma) = \sum_{U \in \gamma} p(U).$$

Safety

- Any instance of global BGP execution is safe with respect to an initial state π_0 if and only if for an activation probability vector $\mathbf{P}$ there is no cyclic state.
- PCTL: $\mathbf{P}_{\geq 1}[\mathbf{GF}\,\pi \rightarrow \mathbf{FG}\,\pi], \;\; \forall \pi \in Q(S).$
- CTL*: $\mathbf{A}[\mathbf{GF}\,\pi \;\rightarrow\; \mathbf{FG}\,\pi], \;\; \forall \pi \in Q(S).$



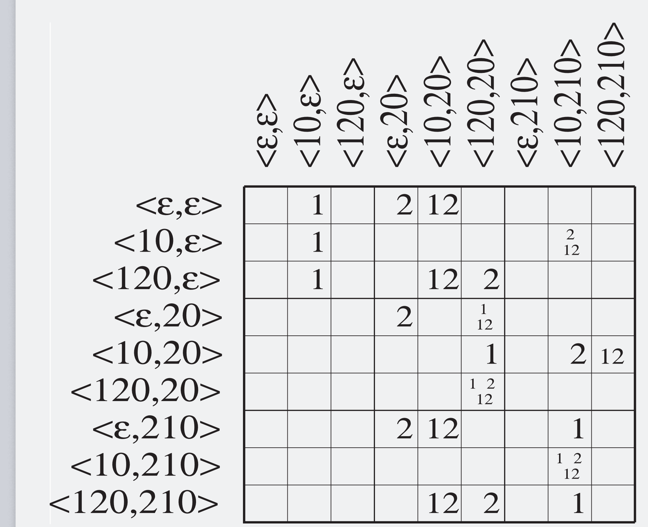
Safety illustration

Convergence Time

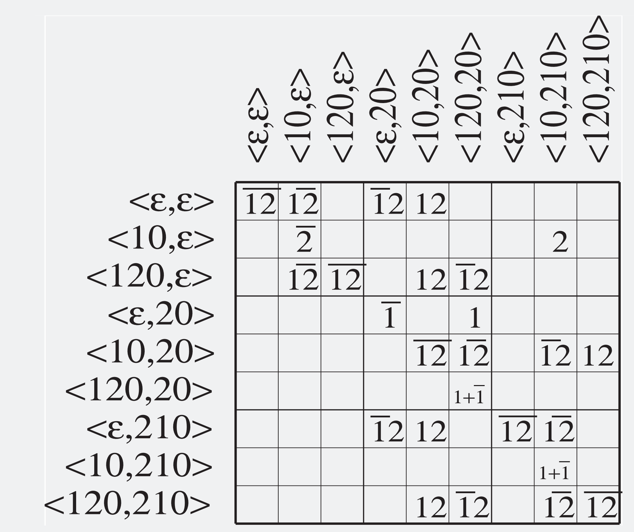
- We define a state reward function $\rho(\pi)$ as: $\rho(\pi) = 1, \;\; \forall \pi \in Q(S)$.
- Let δ denote a unique absorbing state. The number of transitions made until δ is reached may be expressed as: $\mathcal{R}_{=?}[\mathbf{F}\,\delta]$.

Example

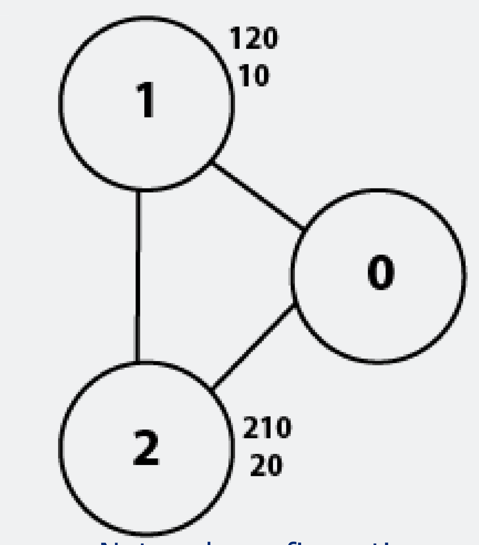
- We used example by Viswanathan et al., and PRISM for model checking.
- Network configuration is deterministically unsafe but probabilistically safe.



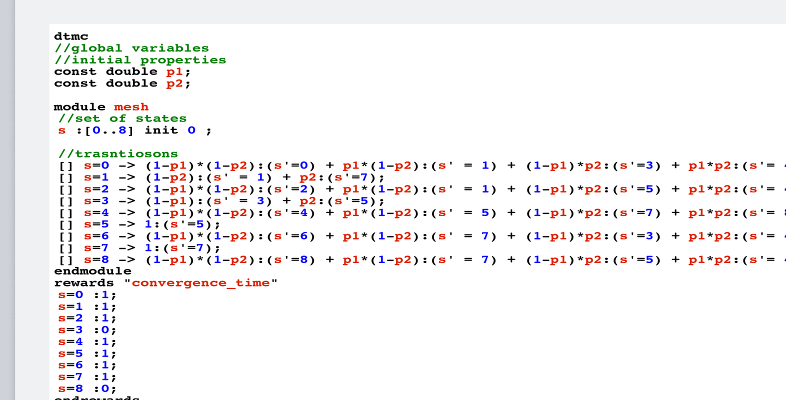
Transition matrix



Probabilistic transition matrix



Network configuration



PRISM model checker code

Safety with p1=p2=1		Safety with p1=p2=0.9	
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 0) \rightarrow (\mathbf{FG}s = 0)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 0) \rightarrow (\mathbf{FG}s = 0)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 1) \rightarrow (\mathbf{FG}s = 1)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 1) \rightarrow (\mathbf{FG}s = 1)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 2) \rightarrow (\mathbf{FG}s = 2)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 2) \rightarrow (\mathbf{FG}s = 2)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 3) \rightarrow (\mathbf{FG}s = 3)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 3) \rightarrow (\mathbf{FG}s = 3)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 4) \rightarrow (\mathbf{FG}s = 4)]$	*FALSE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 4) \rightarrow (\mathbf{FG}s = 4)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 5) \rightarrow (\mathbf{FG}s = 5)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 5) \rightarrow (\mathbf{FG}s = 5)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 6) \rightarrow (\mathbf{FG}s = 6)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 6) \rightarrow (\mathbf{FG}s = 6)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 7) \rightarrow (\mathbf{FG}s = 7)]$	*TRUE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 7) \rightarrow (\mathbf{FG}s = 7)]$	*TRUE*
$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 8) \rightarrow (\mathbf{FG}s = 8)]$	*FALSE*	$\mathbf{P}_{\geq 1}[(\mathbf{GF}s = 8) \rightarrow (\mathbf{FG}s = 8)]$	*TRUE*

- The calculated convergence time is infinity.
- Case p1=p2=1: the policy is not safe. Hence, the convergence time is infinity.
- Case p1=p2=0.9: the policy is safe. However, the absorbing state is not unique and the model checking total calculated reward is infinity.

References

[1] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," *IETF, RFC 1771*, March 1995.
[2] E. M. Clarke and E. A. Emerson, "Design and synthesis of synchro- nization skeletons using branching-time temporal logic," in *Logic of Programs Workshop*, London, UK, 1982, pp. 52–71.
[3] R. Viswanathan, K. K. Sabnani, R. J. Holt, and A. N. Netravali, "Expected convergence properties of BGP," *Computer Networks*, vol. 55, no. 8, pp. 1957–1981, June 2011.
[4] Probabilistic Symbolic Model Checker [Online]. Available: http://www.prismmodelchecker.org/.

SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

ICNP 2011
19th International Conference on Network Protocols
October 2011, Vancouver, Canada

Communication Networks Laboratory
http://www.ensc.sfu.ca/~ljilja/cnl