# Analysis of Weather Forecasting Model in PRISM

Asad Ahmed, Adnan Rashid and Sohail Iqbal School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan. Email: [14dphditaahmed, 14dphditarashid, sohail.iqbal] @seecs.nust.edu.pk

Abstract-Weather forecasting provides an important information for general public. Analysis of weather forecasting models, through traditional simulation techniques, are not exhaustive and thus not accurate. Probabilistic model checking, a formal methods technique, as a complementary approach provides an accurate analysis for probabilistic models. In this paper, we have analyzed a simple probabilistic weather forecasting model for Islamabad weather using PRISM model checker. The reachability and prediction capability of model has been analyzed using quantitative and qualitative properties in PRISM. Analysis shows that all the states of weather forecasting model are reachable and model is biased. In addition, we have also implemented this model in MATLAB and compared the results of both implementation, quantitatively. We have also presented a comparison between PRISM and MATLAB implementation of weather forecasting model, qualitatively.

#### I. Introduction

Weather has an important effect on the daily life routine. A weather forecast can be beneficial to administrators of Electrical company, local road authority and even for a student: in deciding whether to carry the umbrella or not. A good material is available on the weather forecast modelling [1].

There are large number of phenomenon in nature which exhibit stochastic behaviour. These stochastic behaviours are modelled as family of random variables sampled at time, i.e.  $X_{t_{n+1}}: t \in T$ , called stochastic processess. Processes are categorized as either Discrete or continuous stochastic processes. For countable t, stochastic process is called discrete stochastic process and for uncountable t it is classified as continuous stochastic process. Markov processes are a sub class of stochastic process, under the assumption, which allows mathematically tractable analysis of process. The assumption, which reduces the process into a Markov process is: conditional cumulative distribution function of  $X_{t+1}: t \in T$  only depends upon previous value of  $X_{t_n}$ , and is independent of all previous values of  $X_{t_0}, X_{t_1}, \dots, X_{t_{n-1}}$ . Markov processes can be further regarded as Discrete and continuous processes. Among the two categories, discrete Markov chains are related to current work. Discrete Markov chains are completely described by the tuple (S, P), under the assumption of timehomogeneity [2], where S is set of finite states and P is matrix of transition probabilities of states.

Thomas and Fiering model [3], made first attempt to model the stream flow as stochastic time series, with the intent to explore some patterns of stream flow. In [4], effects of data on such techniques of modelling have been investigated.

Tools used for analyzing the real world problems in the scientific community are based on simulations, which are

inherently not exhaustive in analyzing the system under consideration. Formal methods [5] are an alternative approach, which use mathematical reasoning, for compensating scalability and inaccuracy issues of simulation. Being exhaustive and sound, formal methods draw the attention of researchers in safety critical applications such as [6], [7] and [8]. Main objectives of using Formal methods are specification and verification of mathematical models of hardware or software [9]. Specification involves description of a system and its properties using defined syntax and semantics. Formal methods such as Z [10], VDM [11], CSP [12], CSS [13] and LOTOS [14] are one of many others used for specification. Verification answers the correctness of specification w.r.t to its implementation. Model checking and theorem proving are methods used for verification methods [9].

Growing use of embedded systems poses challenge of more reliable and accurate analysis requirement in early stages of design [15]. Model checking, is a technique, relies on a finite model of a system and checks, exhaustively, the system for given property. It performs an automated verification and provides a counter example if property is not satisfied. Model checkers such as PRISM [16], [17], YMER [18] and VESTA [19] have widely been accepted as routine practice in different fields [9]. In this paper, as per requirement of problem, probabilistic model checking [15] has been applied to verify the properties of mathematical model of weather forecast, i.e. biasing and predictability of model. It constitutes an automated verification of probabilistic or stochastic processes, usually a variant of Markov process. It offers systematic and exhaustive analysis of model of system under consideration by finding the answers to questions of correctness and reliability, quantitatively as well as qualitatively. Property specification for the dynamic systems is the motivation for Temporal Language. Temporal language is based upon either Computational Tree Logic (CTL) or Propositional Linear Temporal Logic (PLTL) [20]. PLTL is fragment of CTL, CTL constitutes X (Next), F (Future) and G (Globally) along with Boolean connectors is used to express properties [21]. Where, PLTL constitutes A (For all) and E(There exist), called path quantifiers, along with Boolean connectors is used to write the properties of system [21]. An extension of CTL is Probabilistic CTL to describe the Discrete Markov Chain properties [22]. Probabilistic model checking, and in particular PRISM model checker has been successfully applied in different fields to find faults, e.g. [23], [24], [25], [26], [27].

In this paper, we are interested in exploring the different aspects of Markov weather forecasting model in PRISM [16] model checker. Weather forecasting model of second order has been considered [28], which considers consecutive days weather inter dependency, only. For this model, transition

probabilities have been calculated by using the data available from <code>www.accuweather.com</code>, by considering precipitate as measure of rainy or sunny. Model is implemented in PRISM for analyzing the correctness as well as predictive power of model. To best of our knowledge, formal verification has never been applied to assess the any kind of wether forecasting models.

In Section II, we present weather forecast model for Islamabad as Markov process. Then, in Section III, implementation in PRISM is described. Section IV presents its implementation in MATLAB and the comparison of the results is also presented and Section V concludes the paper.

#### II. WEATHER FORECASTING MODEL

Markov processes are used to model variety of natural phenomenon [29]. A discrete Markov model relies on the memory less property, which equips the resulting time series with mathematical tractability. Additionally, history of phenomenon is incorporated by defining the number of state variables. Number of state variables determines the order of Markov chain. But, introduction of variables in the state results in exponential growth of parameters to be estimated. It results in decreasing the predictive power of model.

#### A. Discrete Markov Chain (DTMC)

A Markov chain consists of a set of states and a set of labeled transitions between the states. A state models a condition of interest and a transition is either probability of transition or rate of transition. In case of probability label, resulting Markov chain is categorized as Discrete-time Markov chain. DTMC evolves step by step,i.e. from state state  $s_n$  to  $s_{n+1}$ , according to one step transition Probabilities matrix of transition is

The n step transition probability matrix can be calculated by computing (n-1) fold multiplication of one step transition probabilit matrix,i.e

$$P^n = P^1 P^{n-1}$$

## B. Islamabad weather as Markov Model

Probability of precipitation has been an important parameter for weather forecast [30]. In our work, we have considered the precipitate as measure of rainy or sunny. Probability of precipitate has been found by considering the weather conditions from November to March at <a href="https://www.accuweather.com">www.accuweather.com</a>.

A second order Markov chain has been used to model the weather of Islamabad. Fig. 1 represents the behavior of the weather of Islamabad. If it sunny on a particular day, it may remain sunny on the next day or it will change to rainy day and so on.

The equivalent second order Markov model for the weather forecasting system is depicted in the Fig. 2. In this model, S



Fig. 1. Different weather conditions of consecutive days

stands for sunny and R stands for rainy, where  $P_{SS}$ ,  $P_{SR}$ ,  $P_{RR}$  and  $P_{RS}$  are the state transition probabilities. State 1 of our model represents the two consecutive sunny days. State 2 models the sunny and rainy weather on two consecutive days and rainy weather on two consecutive days is represented by state 3. State 4 models the rainy and the sunny weather on two consecutive days.

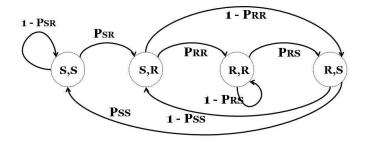


Fig. 2. Markov Model for proposed weather model

For a discrete Markov model, the sum of probabilities from a given state is equal to 1. Our model is consistent with this property.

# C. Transition/Markov Matrix

Transition probabilities of discrete Markov chain captures all information that is required for given problem. The number of states in a Markov model defines the order of transition probability matrix. As we are having four states for our equivalent Markov model of the underlying system, we have a transition matrix of order four which contains the transition probabilities of the respective states. The Markov matrix is given below:

$$P = \begin{bmatrix} 1 - P_{SR} & P_{SR} & 0 & 0\\ 0 & 0 & P_{RR} & 1 - P_{RR}\\ 0 & 0 & 1 - P_{RS} & P_{RS}\\ P_{SS} & 1 - P_{SS} & 0 & 0 \end{bmatrix}$$

#### D. Calculation of Transition Probabilities

We have calculated the transition matrix from the data on website www.accuweather.com. We have taken the precipitate measure of all the data, then defined sunny for precipitate value 0 and rainy for else. For example, for one month duration, we had sample space of 29 by considering the consecutive days as sample. From the data, occurrence of two consecutive sunny day event was calculated and then applied empirical probability formula.

The empirical probability of an event is defined as the ratio of total number of outcomes of that event to the total number of trials. If n is the total number of trials and k is the number

of outcome of an event then empirical probability is given by formula below:

$$Empirical\ Probability = k/n \tag{1}$$

The data snapshot from www.accuweather.com is shown in Fig. 3.

Pakistan WEATHER	Islamaba LOCAL WE	ATHER	•		
Now V	Veekend	Extend	ied	Month	Satellite
< February 20	14	View:		Marc	ch v 2014
	High	Low	e	precip	Snow
Sat 3/1/2014	13°	90		0 mm	0 CM
Sun 3/2/2014	21°	139	•	0 mm	0 CM
Mon 3/3/2014	140	119	•	35 mm	0 CM
Tue 3/4/2014	19°	10	,	2 mm	0 CM
Wed 3/5/2014	21°	90		0 mm	0 CM
Thu 3/6/2014	23°	90		0 mm	0 CM
Fri 3/7/2014	21°	90		0 mm	0 CM
Sat 3/8/2014	26°	90		0 mm	0 CM
Sun 3/9/2014	19°	129	,	13 mm	0 CM

Fig. 3. Data Snapshot from accuweather 1

In Fig. 3, in the precipitation column, if it is 0mm, we have assumed it as a sunny day and for any precipitation greater than 0mm, we have taken it as a rainy day. Data is collected for past 6 months and from that we have calculated the transition probabilities of the underlying state space.

### III. IMPLEMENTATION IN PRISM

We have implemented the underlying model in PRISM. PRISM offers the analysis by verifying quantitatively as well as qualitatively. Former furnishes us with the soundness of result computed over all possible states of model, while later ensures the expected behavior or anomalous behavior, if any.

## A. Transient Probabilities

The transient behavior is analyzed for n = 1, 2, ..., 7, this reveals the dynamics of our model. As, we go from n = 1 to

n=7 we find some states reachable and others not or less probable, comparatively.

For n = 1, table shows that some states are not reachable given the initial states which is in accordance with our model.

For n=2,3,...,7, we observe that all states are reachable as well as the probability of sunny day is more than others, comparatively. This biased prediction is due to the sampling of data. Six months of data from November to March was used, which is not rainy season in Islamabad.

TABLE I. PROBABILITIES OF TRANSITIONS OF STATES AFTER ONE TIME UNIT (1 DAY)

Transition Pro	Transition Probability		Final State				
Transition Tre			SR	RR	RS		
	SS	0.8333	0.1667	0	0		
Initial State	SR	0	0	0.2	0.8		
Initial State	RR	0	0	0.834	0.166		
	RS	0.467	0.533	0	0		

TABLE II. PROBABILITIES OF TRANSITIONS OF STATES AFTER TWO TIME UNIT (2 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.6945	0.1389	0.0333	0.1333	
	SR	0.3733	0.4267	0.1669	0.0333	
	RR	0.0778	0.0889	0.6944	0.1389	
	RS	0.3889	0.0778	0.167	0.4267	

TABLE III. PROBABILITIES OF TRANSITIONS OF STATES AFTER THREE TIME UNIT (3 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.6409	0.1868	0.0556	0.1167	
	SR	0.3266	0.08	0.2242	0.3691	
	RR	0.1296	0.0870	0.5965	0.1868	
	RS	0.5232	0.0.2928	0.1044	0.08	

TABLE IV. PROBABILITIES OF TRANSITIONS OF STATES AFTER FOUR TIME UNITS (4 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.5885	0.1690	0.0837	0.1587	
	SR	0.4444	0.2513	0.2028	0.1014	
	RR	0.1952	0.1212	0.5145	0.1690	
	RS	0.4733	0.1298	0.1455	0.2513	

# B. Quantitative Analysis

Here are some of the quantitative properties.

<sup>1</sup>http:// www.accuweather.com

TABLE V. PROBABILITIES OF TRANSITIONS OF STATES AFTER FIVE TIME UNITS (5 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.5645	0.1827	0.1035	0.1492	
	SR	0.4177	0.1281	0.2193	0.2348	
	RR	0.2416	0.1227	0.4529	0.1827	
	RS	0.5117	0.2129	0.1472	0.1281	

TABLE VI. PROBABILITIES OF TRANSITIONS OF STATES AFTER SIX TIME UNITS (6 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.54	0.1736	0.1228	0.1634	
	SR	0.4577	0.1948	0.2083	0.1390	
	RR	0.2865	0.1377	0.4020	0.1736	
	RS	0.4862	0.1536	0.1653	0.1948	

TABLE VII. PROBABILITIES OF TRANSITIONS OF STATES AFTER SEVEN TIME UNITS (7 DAYS)

Transition Probability		Final State				
		SS	SR	RR	RS	
Initial State	SS	0.5263	0.1771	0.1370	0.1593	
	SR	0.4463	0.1504	0.2126	0.1906	
	RR	0.3198	0.1403	0.3625	0.1771	
	RS	0.4961	0.1849	0.1684	0.1504	

If you are currently at any of the state, i.e. State SS, SR, RR or RS. The probability that state RS will occur in future is specified by the following quantitative property.

$$P = ?[F \ x = 3]$$
 (2)

where in eq.(2), P = ? means: what is the probability of occurrence of a state ?, F stands for Future and  $[F \ x = 3]$  means state RS will occur in future. In our state space model, this probability measure resulted in to 1.

If you are currently at any of the state, i.e. State SS, SR, RR or RS. The probability that state RR will occur in future is specified by the following quantitative property.

$$P = ?[F \ x = 2]$$
 (3)

If you are currently at any of the state, i.e. State SS, SR, RR or RS. The probability that state SR will occur in future is specified by the following quantitative property.

$$P = ?[F \ x = 1] \tag{4}$$

If you are currently at any of the state, i.e. State SS, SR, RR or RS. The probability that state SS will occur in future is specified by the following quantitative property.

$$P = ?[F \mid x = 0] \tag{5}$$

The teachability of all states is analysed by using the quantitative measure of probability of all states in future.

If you are currently at state SS. The maximum probability that state RS will occur after 1 time unit (1 day) is specified by the following quantitative property and it will be zero as can be seen from Markov model.

$$P_{max} = ?[F <= 1 \ x = 3] \tag{6}$$

where in eq.(6),  $P_{max}=?$  means what is the maximum probability of occurrence of a state and  $[F<=1\ x=3]$  means state RS will occur after 1 day in future.

## C. Qualitative Analysis

Here are some of the qualitative properties.

If you are currently at State SS. The probability that you would remain in that state for next eight days is always greater than 0.833 and is specified by the following qualitative property.

$$P > 0.8333 [F < 8 x = 0]$$
 (7)

where in eq. (7), P stands for Probability of a state occurrence and  $[F < 8 \ x = 0]$  means state SS will occur in future within first 8 days.

If you are currently at State RS. The probability that you would remain in that state for next eight days is always greater than 0.833 and is specified by the following qualitative property.

$$P > 0.8333 [F < 8 \ x = 2]$$
 (8)

The above results show the biased nature of estimates, which is logical as the transition probabilities shows that the probabilities of the states are higher then others.

## IV. IMPLEMENTATION IN MATLAB

Given system is also implemented in MATLAB. To implement the system we need, a vector containing probabilities of states and transition probabilities matrix.

$$\Pi^{(0)} = [\pi_0^{(0)}, \pi_1^{(0)}, \pi_2^{(0)}, \pi_3^{(0)}] \tag{9}$$

Eq.(9) is a vector of initial probabilities of state, i.e. in our case SS, SR, RR, RS.

The time evolution of Markov chain and its transient analysis governing equation is given below:

$$\Pi^{(n)} = \Pi^{(0)} \mathbf{P}^n \tag{10}$$

Where,  $\mathbf{P}$  in the transition/Markov matrix as given in section II-A.

This equation is used to plot the probabilities at nth time step, i.e. nth day from current state. Consecutive sunny days state is considered and probabilities of next six days are plotted. Graph in Fig. 4 shows that as the model evolves in time, probabilities of making a transition to next state increases and probability of state being in initial state decreases.

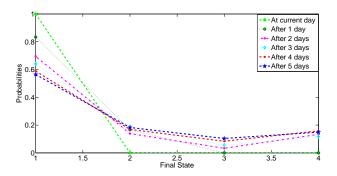


Fig. 4. Transition Probabilities of the States

The comparison between and MATLAB results is presented in the table given below:

TABLE VIII. A COMPARISON B/W PRISM AND MATLAB RESULTS

PRISM							
Transition Probability		Final State					
Transition F	Tobability	SS	SR	RR	RS		
	1 day	0.833	0.1667	0	0		
	2 days	0.6945	0.1389	0.0333	0.1333		
	3 days	0.6409	0.1868	0.0556	0.1167		
After days	4 days	0.5885	0.1690	0.0837	0.1587		
	5 days	0.5645	0.1827	0.1035	0.1492		
	6 days	0.5400	0.1736	0.1228	0.1634		
	7 days	0.5263	0.1771	0.1370	0.1593		
		MAT	LAB				
Transition F	Probability	Final State					
Transition F	Tobability	SS	SR	RR	RS		
	1 day	0.833	0.1667	0	0		
	2 days	0.6945	0.1389	0.0333	0.1333		
	3 days	0.6409	0.1868	0.0556	0.1167		
After days	4 days	0.5885	0.1690	0.0837	0.1587		
	5 days	0.5645	0.1827	0.1035	0.1492		
	6 days	0.5400	0.1736	0.1228	0.1634		
	7 days	0.5263	0.1771	0.1370	0.1593		

Table VIII, presents the results implemented in MATLAB and PRISM. Conditions for both environments are same and

results are compared. Initial state of weather is SS and time step is varied from 1 to 7. Results of PRISM compliment the results of MATLAB.

### V. CONCLUSION

In this paper, Markov model of weather of Islamabad, for the given length of data, is implemented in PRISM. From transient behavior and qualitative properties, it is found that model is biased for sunny weather. From quantitative properties for model, it is found that probability of reaching any state is not zero which shows the correctness of model, i.e. no dead lock state occurs over the period of time. Model is also implemented in MATLAB for quantitative measure of probabilities, which are found consistent with PRISM. Moreover, model can be refined to tune the weather forecasting by considering the data size, order of Markov model and state variables. Property specification is another open area to work out, which allows verifying the specific behavior of model.

#### REFERENCES

- [1] Ian T Jolliffe and David B Stephenson. Forecast verification: a practitioner's guide in atmospheric science. John Wiley & Sons, 2012.
- [2] Gunter Bolch, Stefan Greiner, Hermann de Meer, and Kishor Shridharbhai Trivedi. Queueing networks and Markov chains: modeling and performance evaluation with computer science applications. John Wiley & Sons, 2006.
- [3] P Jaya Rami Reddy. A text book of hydrology. Firewall Media, 2005.
- [4] Daniel S Wilks and Robert L Wilby. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23(3):329–357, 1999.
- [5] J.-R. Abrial. Faultless systems: Yes we can! Computer, 42(9):30–36, Sept 2009.
- [6] Binyameen Farooq, Osman Hasan, and Sohail Iqbal. Formal kinematic analysis of the two-link planar manipulator. In Formal Methods and Software Engineering, pages 347–362. Springer, 2013.
- [7] Robert S Boyer, Matt Kaufmann, and J Strother Moore. The boyer-moore theorem prover and its interactive enhancement. *Computers & Mathematics with Applications*, 29(2):27–62, 1995.
- [8] Wael M Elseaidy, John W Baugh Jr, and Rance Cleaveland. Verification of an active control system using temporal process algebra. *Engineering* with Computers, 12(1):46–61, 1996.
- [9] Edmund M Clarke and Jeannette M Wing. Formal methods: State of the art and future directions. ACM Computing Surveys (CSUR), 28(4):626– 643, 1996.
- [10] J Michael Spivey and JR Abrial. The Z notation. Prentice Hall Hemel Hempstead, 1992.
- [11] Cliff B Jones. *Systematic software development using VDM*, volume 2. Prentice-Hall Englewood Cliffs, NJ, 1986.
- [12] Uwe Wolter. Csp, partial automata, and coalgebras. Theoretical Computer Science, 280(1):3–34, 2002.
- [13] Xiao Jing, Wu Guo-qing, and Chen Shu. A method for transforming workflow processes to css. In Computer Engineering and Technology, 2009. ICCET'09. International Conference on, volume 1, pages 295– 299. IEEE, 2009.
- [14] Tommaso Bolognesi and Ed Brinksma. Introduction to the iso specification language lotos. *Computer Networks and ISDN systems*, 14(1):25– 59, 1987.
- [15] Christel Baier, Joost-Pieter Katoen, et al. Principles of model checking, volume 26202649. MIT press Cambridge, 2008.
- [16] PRISM. http://www.prismmodelchecker.org/, 2014.
- [17] T Nipkow et al. Advances in probabilistic model checking. Software Safety and Security: Tools for Analysis and Verification, 33:126, 2012.
- [18] Håkan LS Younes. Ymer: A statistical model checker. In Computer Aided Verification, pages 429–433. Springer, 2005.

- [19] Axel Legay and Mahesh Viswanathan. Simulation+ hypothesis testing for model checking of probabilistic systems. In QEST, page 3, 2009.
- [20] E Allen Emerson and Joseph Y Halpern. sometimes? and not never? revisited: on branching versus linear time temporal logic. *Journal of the ACM (JACM)*, 33(1):151–178, 1986.
- [21] Béatrice Bérard, Michel Bidoit, Alain Finkel, François Laroussinie, Antoine Petit, Laure Petrucci, and Philippe Schnoebelen. Systems and software verification: model-checking techniques and tools. Springer Publishing Company, Incorporated, 2010.
- [22] Marta Kwiatkowska. Model checking for probability and time: from theory to practice. In Logic in Computer Science, 2003. Proceedings. 18th Annual IEEE Symposium on, pages 351–360. IEEE, 2003.
- [23] T. Herman. Probabilistic self-stabilization. Information Processing Letters, 35(2):63–67, 1990.
- [24] G. Norman and V. Shmatikov. Analysis of probabilistic contract signing. *Journal of Computer Security*, 14(6):561–589, 2006.
- [25] M. Kwiatkowska, G. Norman, J. Sproston, and F. Wang. Symbolic model checking for probabilistic timed automata. In Y. Lakhnech and S. Yovine, editors, Proc. Joint Conference on Formal Modelling and Analysis of Timed Systems and Formal Techniques in Real-Time and Fault Tolerant Systems (FORMATS/FTRTFT'04), volume 3253 of LNCS, pages 293–308. Springer, 2004.
- [26] M. Kwiatkowska, G. Norman, D. Parker, and M.G. Vigliotti. Probabilistic mobile ambients. *Theoretical Computer Science*, 410(12–13):1272–1303, 2009.
- [27] Frits Dannenberg, Marta Kwiatkowska, Chris Thachuk, and Andrew Turberfield. DNA walker circuits: computational potential, design, and verification. In D. Soloveichik and B. Yurke, editors, *Proc. 19th International Conference on DNA Computing and Molecular Programming* (DNA 19), volume 8141 of LNCS, pages 31–45. Springer, 2013.
- [28] Ann Henderson-Sellers and Kendal McGuffie. The Future of the World's Climate: A Modelling Perspective. Elsevier, 2011.
- [29] Mark Pinsky and Samuel Karlin. An introduction to stochastic modeling. Academic press, 2010.
- [30] Eric Bickel and Seong Kim. Verification of the weather channel probability of precipitation forecasts. *Monthly Weather Review*, 136(12):4867–4881, 2008.