CHAPTER II

LITERATURE SURVEY

2.1 INTRODUCTION

This chapter provides a review on different handover algorithms existing in the literature. The research work related to handover initiation, handover target network selection and handover execution in coordination with cross layer is explained. After reviewing the related work, research gaps are outlined in the last section.

Nowadays, wireless technologies have been experiencing huge advancements to deliver a variety of applications with guaranteed QoS while moving across a diverse range of mobile access networks. The wireless networks, consisting heterogeneous networks, differ technologically and no individual network can satisfy end users in all respects including low cost, high bandwidth and maximum coverage area. Hence it is obligatory to have integration and coordination of different kinds of networks in upcoming wireless networks. To enjoy service continuity along with multimedia supported applications and to promise expected QoS, efficient handover algorithms need to be incorporated.

2.2 HANDOVER INITIATION ALGORITHMS

A comprehensive survey of vertical mobility management schemes in wireless networks is presented in [27, 28]. Various design challenges such as QoS classes, resource allocation for supporting service differentiation, pricing are highlighted and a QoS based context aware handover mechanism is proposed in [27]. An overview about handover initiation control techniques is given in [29] which describes different approaches and initiation techniques based on radio link measurements to detect handover requirement. In literature, various handover initiation algorithms based on single or multiple criteria such as signal strength, BER, distance and data rate have been implemented. Based on link quality, a number of methods have been proposed to estimate handover initiation time [30−36]. The authors in [30] have made efforts to

Bit Error Rate(误比特率)

minimize average handover delay and handover rate by considering BER in conjunction with signal strength as criteria for handover decision. The authors in [31] and [32] consider interference and data rate along with RSS to take decision about handover initiation. In order to perform handover near boundary area, the distance between MN and BS is also to be considered as an important factor for handover initiation. The number of handovers can be reduced significantly by selecting appropriate handover parameters including threshold signal strength and Hysteresis margin [33]. The hysteresis margin (difference between RSS of serving BS and TBS) and absolute threshold for signal strength, are determined to optimize the handover performance in [34]. The authors in [35] and [36] have utilized the location of MN to find optimum hysteresis margin. The threshold signal strength is predefined in most of the link quality based handover algorithms. However, it may cause ambiguity in the handover decision as the handover conditions like channel parameters, user velocity and user preferences vary with time. The adaptive signal strength has been utilized in [37] to present an application specific signal strength tuning mechanism in which wireless channel variation, network layer latency and application QoS demands are taken into account for improving the system resource utilization. Some researchers have shown the impact of mobile user's velocity on handover performance in wireless networks. A velocity adaptive handover is proposed in [38] in which the threshold RSS varies according to the velocity. However, it has been proved to be effective for higher mobility scenario only and there is no improvement for low velocity users. The authors in [39] designed and implemented domain based approach referred to as Handoff Aware Wireless Access Internet Infrastructure (HAWAII). It uses specified path setup schemes which install host based forwarding entries in specific routers to support intra-domain micro-mobility. With the help of the movement prediction achieved by linear regression model and keeping track of the signal strength of mobile users, Layer−3 (L3) handover activities can be executed prior to Layer−2 (L2) handover. Such a scheme is proposed by [40] to reduce total handover latency by Δt (prediction interval) in mobile WiMAX. A generic link layer technique making use of upper layers to assist handover protocols is used in [18]. However, it does not specify any particular mechanism for obtaining the link layer triggers. The reference [41] gives a performance comparison of various handover mechanisms in WiMAX referred to as: the Link-layer handover algorithm [42], Pre−coordination handover scheme [43], Passport handover scheme [44] and FINCH (Fast Intra Network and Cross layer Handover) [45]. Out of these mechanisms, FINCH is found to be having the least handover delay. An adaptive RSS threshold using estimation of boundary area is used to initiate handover in time [46]. The estimated distance gives the distance remaining to reach handover point which in turn estimates the probability of handover failure.

2.3 CROSS LAYER BASED HANDOVER DECISION ALGORITHMS

Because of the numerous advantages of the cross layer approach, these techniques are adopted by various standardization bodies like IEEE, IETF, and 3GPP. Likewise, most of the recent standards including IEEE 802.11 WLANs, IEEE 802.16 WMANs and IEEE 802.21 are also adopting cross layer design to refine the existing handover procedure [47].

One resource consuming task in handover management is to collect and process information extracted from event sources of different layers into a standardized trigger format. These triggers can further be used to take a handover decision in order to enable seamless handover. The handover performance is greatly affected by the time when a handover trigger is generated [48, 49]. The authors in [48] have investigated the impact of L2 triggering time on the signaling cost, packet delivery cost, total overhead cost, and buffer space. The design parameters, which may affect the handover trigger time, have been tuned to reduce the outage probability in [50].

Assuming the ability of MN to support multiple interfaces, the MIH facilitates fast handover by exchanging event information among entities from different technologies through cross layering [51]. The authors in [51] have shown the support of IEEE 802.21 functionalities to attain seamless mobility between IEEE 802.11 and IEEE 802.16. By making use of facilities provided by the IEEE 802.21 standard, an integration of broadcast technologies like DVB−H in the wireless heterogeneous environment is also possible [52].

The authors in [53] propose a dual link based handoff approach in which handoff latency and handoff trigger threshold are predicted based upon user history and the pattern of MN movement. The dual link handoff allows MN one link to connect with the new subnet while keeping another link connecting with the current one. In [54], layer agents are defined at each layer to capture the necessary parameter and report to Link Information Manager (LIM). The LIM forwards the parameter to a Decision Engine (DE) which notifies a cross layer manager about an impending handover. The cross layer manager intimates higher layers about handover to the indicated interface. According to [19], instead of exporting information between layers, it is more useful to import information into a separate management database visible across all layers. In addition, threshold RSS is suggested to be adaptive for inter and intra handovers but it has not been implemented in this paper. To achieve this goal, the authors propose a bow−tie architecture with a well defined cross layer service interface at the centre to provide a unifying link abstraction to the five layer protocol stack. The mobility management forms the right side of the bow−tie which is served by a cross layer database visible to all layers. In [55], the collection and processing of cross layer information is performed with the help of cross layer trigger manager in which triggers are generated by event sources/trigger producers. This framework enables handover decision making through information from both MN's local stack and the network. A module named TRG is introduced to collect cross layer information from different event sources, e.g. a WLAN interface, mobility management software or an adaptive application. The TRG translates the events into triggers and delivers to their consumers after processing. Trigger management for collecting and processing the cross layer information is also shown in [26] but it has nowhere described any predefined or adaptive threshold for trigger producers or event sources. A modular architecture is employed in [56] to monitor interfaces to different technologies including Ethernet, IEEE 802.11b and GPRS. An event handler running in user space is used to manage events, read from an Event Queue, where events are inserted by modules monitoring the network interfaces. A trigger Management architecture for a vertical handover controller system has been presented in [57] which supports application of different triggering policies and integrates them to feed Hierarchical IP , MIP, and Session handovers. The handover decision is made based upon predefined rules or parameters including cost, power consumption, and wireless QoS support. Each interface gets calculated preference value according to priorities of different parameters. Trigger Manager may also use different policies at different times or when in different contexts, without having to regenerate rules. Also triggers can be formed depending upon the system or application context instead of mobility management procedures. Effective and timely link−layer trigger mechanisms can significantly improve the handover performance. If the handover initiation is on appropriate time, handover latency and service disruption time is reduced as well as preceding interface is utilized at the maximum. S.Yoo et al. [58] have implemented timely fired link triggers to reduce handover signaling delay in heterogeneous networks. The authors have generated the link triggers using predicted signal strength and neighbor network information obtained through MIHF. The handover latency from has been reduced from 450ms to 55ms. The accurate network conditions are provided timely by the link layer triggers that are fired at MAC sub layer. But the RSS of target network is mentioned nowhere to verify the threshold condition. Chi Ma, et al. [59] have proposed a velocity optimized handover mechanism for WiMAX networks by using handover failure probability value to generate LGD trigger. The distance travelled by user in the target network area is predicted analytically in [60] and is used subsequently to find necessity of a handover to avoid unnecessary handovers. Cheng-Shong Wu et al. [61] have suggested periodic trigger in addition to RSSI trigger for handover in next generation WiMAX femtocells/macrocell overlay environment. This algorithm forces the macrocell users to handover into available femtocells thus avoids unnecessary handover in case of high velocity users. In [62], the prediction triggers are generated using auto-regression integrated moving average model. The RSS values are predicted 1, 5 and 10 steps ahead and compared with real values of RSS. It has been observed that the accuracy is degraded with increase in number of steps. A comparison of user mobility pattern prediction algorithms is performed in [63] in order to increase handover trigger accuracy. The movement pattern of mobile user across multiple cells is analyzed to reduce prediction error. The sensitivity of overall prediction accuracy is found to be dependent upon the configuration of pattern detection algorithms. The prediction error has been reduced up to 45% by this method. In [64], sensitivity analysis of handover margin and time to trigger is carried out for different system load levels and user speeds in LTE network. A fuzzy logic controller is designed to modify the handover margin for optimization of handover. A tradeoff between call drop and signaling overload is achieved. A mobile-receiver centric loosely coupled cross layer design is developed by authors in [65] which is compatible with Wireless Access Protocol (WAP) and robust in the absence of perfect cross layer information. Two proactive schemes, referred to as RTT inflation and RTT equalization, are proposed. The former inflates re-transmit timeout time to suppress premature timeout while the latter prevents false fast re-transmit by equalizing the round-trip delay experienced by all packets. The authors in [66] have proposed cross layer architecture for uninterrupted communication services between a satellite segment and a number of Wi-Fi hot spots. The behavior of TCP is analyzed taking into account the packet loss and delay introduced by MIP during handover. A context aware mechanism is implemented in [67] to find handover preparation time using historical context of handover preparation time. The number of handovers is also reduced by this method. It is shown in [68] that the signaling delay of handoff depends on the link quality, traffic level and the distance between user and its home network. Mohanty [69] presents an LGD trigger algorithm to select handover trigger time for reducing link failure probability for different values of handover delay. The handover signaling delay prediction technique used in this paper is simple, but introduces extra signaling overhead to the system which may not be acceptable for a particular deployment scenario. Moreover, the recommended protocol, HMIP is not suitable for high velocity applications. The authors in [70] introduce a Planned Multicast Group (PMG) approach to strategically arrange all the mesh routers to form a scalable wireless mesh backbone which can facilitate efficient inter gateway handoff supports from multiple layers. The information obtained from L2 and L3, such as the link quality of the new channel and the IP address of the new access point, is utilized to predict the L3 and L5 handoffs in advance so that part of the handoff procedure can be carried out in parallel before the MN completes L2 handoff. However, it leads to enhanced PMG overhead messages generated during handover process. The authors in [71] suggest cross layer fast handover for SIP, which performs IP address reservation and SIP session update by link layer information. A method to reduce probability of false handover initiation with the help of GPS (Global Postioning System) is recommended by the authors in [72]. None of the above mentioned algorithms consider different network overlapping regions. The authors in [73] use MN's trajectory to reduce the false handover in overlapping cells. In [74], an auto regression model is used to select handover trigger time in varying network overlapping environment. However, the LU time is determined based on predefined RSS threshold of TBS, which may result in wastage of network resources in highly overlapping environment. Also the prediction initiation RSS is fixed which may cause too early or too late handover initiation for slow and fast users respectively. The authors in [17] have formulated network resource allocation as a utility maximization problem with rate constraints at network layer and schedule ability constraint at the link layer. Also a distributed sub gradient algorithm is obtained by decomposing the system problem into congestion control, routing and scheduling sub problems for time varying channels and adaptive multi−rate devices. The MIH triggers are generated to initiate handover, using QoS experienced by users for a particular application [75], resulting into improved handover time, end to end delay and throughput. The framework designed in [76] is able to reduce handover delay by utilizing Proxy MIPv6, 802.21 and cross layer design. It reduces mobility related signaling round trip time delay by implementing a micro−mobility protocol at the network side, thus reducing the handover delay associated with network discovery, authentication, address configuration and Duplicate Address Detection (DAD) process.

2.4 FUZZY LOGIC BASED HANDOVER DECISION

The researchers have shown great interest in fuzzy logic while dealing with complex problems associated with handover decision [77, 78, 79, 80]. In [77], three criteria, namely, received power levels, user population and used bandwidth of each BS are used to take a handover decision. The authors in [78] simplify the problem by taking RSS, speed and distance as the inputs to fuzzy logic controller. A multi criteria vertical handoff decision algorithm is designed and implemented in [79]. It decides for handoff using fuzzy values of RSSI, bandwidth, network coverage and preferred network as input to fuzzy controller. The authors in [80] have introduced fuzzy logic for wireless networks to solve complicated problems. Another fuzzy based modular decision system is designed in [81] to reduce handover execution time. The handover parameters are optimized in [82] according to the current traffic and load conditions using a fuzzy controlling technique leading to capacity improvement of up to 9%. The handover requirement estimation is implemented using fuzzy values of RSS, QoS, BER and network coverage in [83]. An immediate or dwell handover can be preferred with the help of fuzzy logic to reduce packet transfer delay [84]. The strength of fuzzy logic has been exploited to take benefits of radio and optical wireless integrated technology. A fair comparison between various MADM techniques referred to as MEW, SAW, TOPSIS and GRA is analyzed in [85]. A fuzzy MADM based numeral solution for vertical handover decision is introduced in [86]. As fuzzy logic has the ability to handle uncertain and conflicting decision metrics, many handover strategies have utilized fuzzy logic in TOPSIS to take decision about target selection during handover [87−89]. It is shown in [90] that the fuzzy membership functions are optimized using a genetic algorithm to select best target network and to reduce the number of handovers. A modified Weighted Product Method (WPM) method is practiced in [91] where Quality of Context (QoC) is utilized to penalize alternatives and weight distribution depends on a fuzzy measure of saliency of context information. Multiple decision makers are used to take a compromised decision in [92]. Ioannis and Drakoulis [93] have followed an energy efficient network selection approach using parameterized utility functions to make it adaptable for different applications. The authors further use fuzzy representation of TOPSIS method to resolve issues of inconsistency related to conflicting decision criteria. The authors in [94] recommend AHP for subjective and entropy method for objective weights to satisfy the QoS requirements of the user. The authors in [95] integrate GRA and AHP for target network selection. The researchers in [96] have proposed a smart decision model to select the best network interface and the best moment to handoff in which the properties of available network interfaces, the system information and the user preferences are utilized to calculate the score function. In [97], an adaptive network fuzzy inference system is investigated where the requirement of human experts can be avoided by incorporating training elements into existing fuzzy handoff algorithms.

2.5 CROSS LAYER BASED HANDOVER EXECUTION ALGORITHMS

There are a number of competing micro−mobility protocols [98] to perform handover, including MIP [99], HMIPv6 [100,101], HAWAII, SIP and Cellular IP (CIP) [102]. An empirical analysis of handoff performance for SIP, MIP and SCTP protocols running over different type of wired and wireless access networks is presented in [103]. MIP yields the highest handoff delay amongst the three mobility protocols MIP, SIP and SCTP. It has been observed that SIP and SCTP yield 33% and 55% lower handoff delay respectively compared to MIP. The performance of handover algorithms can be measured in terms of various parameters like handover delay or latency, bandwidth and resource management, power consumption, handover cost, packet loss and QoS. Advanced handover schemes adopted in IEEE 802.16m and LTE networks to fulfill QoS requirements are discussed in [103].

A large number of cross layer designs have been proposed for handover execution in the literature [104−107]. MIPv6 [108,109] and SIP are the most widely deployed protocols for handling mobility and for session control, respectively but inefficiencies stem from SIP/SDP's unawareness of L3 mobility. The authors in [105] and [106] suggest a cross layer scheme to minimize end to end handover delay by integrating SIP and MIPv6 techniques. Here, CoA is used directly in some messages (short−lived message transactions in call initiation) in [105] but MIPV6 is used which is less efficient than FMIPV6. A cross module is introduced between the Home Agent (HA) and serving−Call Session Control functions (S−CSCF), a database function of control plane of IP Multimedia Subsystem (IMS) [106]. Through the cross−layer module, the binding update message received by HA is translated and forwarded to S−CSCF to trigger the registration procedure in IMS. The cross layer information can also be exploited to optimize the binding update mechanism [107]. FMIPv6 protocol uses the L2 trigger information to solve the unacceptable handover latency and the data loss problems of MIPv6 [110]. The authors in [111] have mentioned an IPv6 based fast handover method for seamless macro−mobility (within different domains) support over WiBro (Wireless Broadband) internetworks. An integration of IEEE 802.16e and FMIPv6 with the incorporation of cross layer events and commands is proposed in [112] to reduce handover latency. The integration of technologies can also be applied in case of handover between 3GPP (UMTS) and 3GPP2 (CDMA2000) networks [113,114]. In this method, Measurement Control L2 message in 3GPP network and Neighbor List L2 message in 3GPP2 network contain the L3 information. Afterwards, router solicitation and advertisement messages are omitted. The authors in [114] have reduced overall service disruption time in heterogeneous networks, but the delay incurred in exchanging messages between ASNGW and HA is not reduced. The handover latency of FMIPv6 handover in mobile WiMAX is reduced in [115] by using a network level protocol called context transfer. It transfers connection information between Access Routers (ARs) and omitted the re−registration and re−entry procedure, but DAD delay, return routability procedure and queuing are ignored in this research work. A similar approach is proposed in [116] to improve handover performance by integrating the 802.16e and FHMIPv6 but the selection mechanism of an appropriate BS is not mentioned. A Fast Intra−Network and Cross layer Handover (FINCH) for intra domain handover is introduced in [117] which eliminates the need of address resolution protocol and forms look up table for IP addressing. In [19], Previous Access Router (PAR) is made to validate the proposed new IP address before MN disconnects from PAR to eliminate DAD delay. The dynamic RSS threshold is calculated using the collective information from handover signaling delay estimation units and speed estimation units present in different layers. A Proxy MIPv6 based solution using the cross layer approach is introduced to utilize backhaul bandwidth and network element’s buffer space [118]. The parallel execution of wireless and wired handover procedures using backhaul link can optimize the handover performance [119]. Yilin Song et al. [110] have figured out the reasons for handover latency and included a smart link layer trigger. The TBS is made to scan a buffering support Bi−binding scheme and the smart link event notification policy in [120]. A link trigger buffer mechanism to reduce delay time for packets in the buffer is suggested in [121]. An experimental study has been conducted in [122] to analyze the impact of MIPv6 mobility on GPRS cellular and 802.11b WLAN. The authors have suggested a number of network layer handover optimization techniques including fast Router Advertisements (RA), RA caching, Binding Update (BU) simulcasting and L3 based soft handovers. A simplified procedure for registration is followed to reduce the amount of signaling along with a quantitative analysis of the algorithm to assess its performance [123]. Another approach termed as the handoff aware TCP adaptation [124], dynamically changes the size of the congestion window and retransmission timeout value during a vertical handoff according to the underlying new network characteristics. It has been envisaged in [125] that link layer detection time can be reduced to three consecutive lost frames because it uses failed frames as the criteria for detection instead of weak signal strength. In case of stations receiving only, beacon interval is identified as the key factor to reduce detection time. The drawback of this algorithm is that its duration increases with the load of the cell and the transmitted frame length. Also the search phase delay is reduced by 20% by using active scanning with its timers MinChannelTime and MaxChannelTime set to 1ms and 10.24ms respectively. An attempt has been made to resolve the latency problem in scanning, L2 execution and re−entry phase to some extent in [126]. Though the authors in [126] have decreased total handover latency, but the latency in handover preparation has been increased comparatively. The authors in [127] have included an interworking layer to resolve a number of handover problems simultaneously and also discussed some open issues for further research. The limitation of [127] is that it considers only downlink traffic and the integration of tightly and loosely coupled cross layer is opted which is very complex.

2.6 CONCLUDING REMARKS

The focus of discussion in this chapter is based on the concept of cross layer to perform handover efficiently and contributions to the existing research area. Many researchers have used cross layer for optimizing handover performance in terms of latency, throughput and packet loss. However, certain research gaps are still present which need to be addressed in order to realize seamless handover in next generation wireless networks.

Most of the above mentioned handover algorithms make use of the standard network mobility management protocol, MIP which is simple to implement but has several shortcomings. MIP latency is composed of handover requirement detection latency and MIP registration. Some of the proposed mobility protocols are able to reduce registration signaling delay but do not address the problem of handover requirement detection delay. Most of the existing handover algorithms implicitly assume that the handover signaling delay is constant, so they give poor performance in real life dynamic environment. The above discussed handover algorithms are not able to give excellent performance in terms of all desired parameters including undisrupted services and network resource utilization. Moreover, existing solutions have focused on one or two phases of handover process such as scanning, L2 handover execution and network re-entry phase while all the phases require attention to reduce overall latency. A few researchers have worked on all steps but have not utilized the cross layering between L2 and L3 to its maximum advantage such as cross layer triggers could be used for proper functioning of fast handover protocols like FMIPv6. In view of the limitations of the existing algorithms, this research aims to design efficient handover schemes which would be able to resolve the mentioned issues to the extent possible.